The Height of Camera: Evaluation of the Effects of Fundamental Feature of 360-degree video on User Experience

Xinru Hu

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
M.Sc. Thesis
Supervisor: Jaakko Hakulinen
April 2018

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
Xinru Hu: The Height of Camera: Evaluation of the Effects of Fundamental Feature of 360-degree video on User Experience
M.Sc. Thesis, 53 pages, 1 index and 12 appendix pages
April 2018

The 360-degree video, aka omnidirectional video, can be used to efficiently create immersive experiences and thus has great potential to be utilized in different contexts and situations, such as museums, working grounds or education. However, new technology always comes with new problems and challenges. Until now, there are still not many research studies related to how basic elements of omnidirectional video influence the user experience. In this thesis, one basic but crucial feature of omnidirectional video, the height of the camera, is studied. A user study was conducted, and feedback was gathered to understand how the camera height influences the user experience under both sitting and standing viewing conditions. The results show that the participants felt more immersion and comfortable when the video height seems close to their actual height. The acceptable heights for sitting and standing are different, standing can offer a better viewing and pleasant experience while sitting tends to offer a safer experience. Hopefully, this thesis can offer some useful information for researchers, designers or other people who are interested in related topics, and encourage them to produce better panoramic solutions for future.

Keywords: 360-degree Video, Omnidirectional camera, Omnidirectional video, Height of Camera, User experience, UX evaluation, Immersion, Presence

Acknowledgements

It’s a great experience to study at the University of Tampere and I would like to give my thanks to those who cared, supported and helped me in these three years.

I would like to express my gratitude to my supervisor Jaakko Hakulinen, for his invaluable advice and help during the process of completing this thesis, as well as his understanding and patience.

I would like to thank Claudia Buchner and Kimmo Ronkainen, who helped me in completing this thesis, especially during the user study session. With their sincere advice and technical support, I could successfully complete my user experiments.

I would also like to express my thanks to Markku Turunen, to all the people in University of Tampere who provided me means to complete my goals and to all the participants who helped me in the user study.

Finally, I want to express my thanks to my parents for their unconditional support, without them I would not be who I am today. Love them.

Tampere, 08.04.2018
Xinru Hu

List of Symbols and Abbreviations

HCI	Human-Computer Interaction
HTI	Human-Technology Interaction
UX	User Experience
VR	Virtual Reality
3D	Three-dimensional
HMD	Head-mounted Display
CAVE	Cave automatic virtual environment
ODV	Omnidirectional Video
CCD	Charge-coupled Device
FOV	Field of View
M	Mean
SD	Standard Deviation
SEM	Standard Error Mean

Contents

1. Introduction 1
2. Omnidirectional Camera and 360-degree Video 4
   2.1 From Ordinary Camera to Omnidirectional Camera 4
      2.1.1 The history of Camera 4
      2.1.2 Evolution of Omnidirectional Camera 5
   2.2 Emergence and Purposes of 360-degree Video 7
   2.3 Previous Work 8
   2.4 Challenges in Viewing 360-degree Video 9
3. User Experience of 360-degree Video 11
   3.1 User Experience and Usability 11
   3.2 Evaluation Methods 13
      3.2.1 Usability Evaluation Methods 13
      3.2.2 User Experience Evaluation Methods 14
   3.3 User Experience Evaluation for 360-degree Video System 15
4. Immersion and Presence 17
   4.1 Definitions of Immersion and Presence 17
      4.1.1 Presence 18
      4.1.2 Immersion 19
   4.2 Previous Work 20
      4.2.1 Applications of Virtual Environments and ODV System 20
      4.2.2 Measurement of Immersion and Presence 20
      4.2.3 Issues that Break Immersion and Presence 21
5. Evaluation Materials and Software 22
   5.1 Videos Used in Experiments 22
   5.2 Software Used in Experiments 25
6. User Study 27
   6.1 Participants 27
   6.2 Procedure and Apparatus 28
   6.3 Data Collection and Analysis 30
7. Results 32
   7.1 Height Selection 32
      7.1.1 Most Natural Height 33
      7.1.2 Most Comfortable Height 34

7.1.3 Highest Acceptable Height 35
7.1.4 Lowest Acceptable Height 37
7.2 User Experience Statements 38
7.2.1 Sitting Condition: Shorter Group 38
7.2.2 Sitting Condition: Taller Group 40
7.2.3 Standing Condition: Shorter Group 41
7.2.4 Standing Condition: Taller Group 42
7.3 Interviews 43
7.3.1 Feedback on the Experiment 44
7.3.2 Influence of Camera Height 44
7.3.3 User Experiences 45
7.3.4 Sitting and Standing 45
7.3.5 Motion Sickness 46
7.4 Presence Questionnaire and Immersive Tendency Questionnaire 46
8. Discussion 48
8.1 Camera Heights 48
8.2 Immersion and Presence 49
8.3 Future Work 50
9. Conclusion 52
References 54
Appendix 58
Appendix A - User experience for ODV application experiment script 58
Appendix B - Script: Statements that will be asked after each single video 62
Appendix C - Consent form 63
Appendix D - Background questionnaire 64
Appendix E - Presence questionnaire 66
Appendix F - Immersive tendency questionnaire 69

1. Introduction

The main idea of this thesis work is to understand while watching 360-degree videos, how the omnidirectional camera height influences the user experience under both sitting and standing viewing conditions. With the development of virtual reality technology, the 360-degree video has shown its unique advantages in various fields, it is necessary to study the experience it brings to people and design better products based on that.

As the main tool for capturing the content of human activities, the camera has always been in the dominant position and has been evolving over time. Traditional cameras can only capture a limited field of view (Kasahara, Nagai, & Rekimoto, 2015), while the recent emergence of the omnidirectional camera, or panoramic camera, brings a revolution for the visual field capturing. Different from the traditional camera, omnidirectional camera covers a 360-degree field of view, in other words, a visual field covers approximately the entire sphere. The 360-degree video, or omnidirectional video (ODV), is the video that has been shot with the omnidirectional camera.

In modern society, interactions are everywhere, it is unavoidable to interact with other (group of) people or the environment. Once the video pioneers have mastered the core technology, ODV will become a powerful tool for individuals or organizations to share stories, experiences and locations (attractions). In recent years, consumers, social media services and different organizations have started to gradually use ODV for various objectives. With mobile devices, viewers can just pan or tilt the phone to change the viewing angle, or, click and drag the screen to freely move the views. Earlier in 2015, Facebook announced that it is going to use ODV to support broadcasts or news feed, to offer immersive video experiences. At about the same time, Google also announced that it is possible to upload and view 360-degree video on YouTube. These ODV contents will bring a new wave and help the media platforms to attract a large number of new audience. Museums, education organizations, air services and many other services or platforms are experimenting with ODV as well. At the same time, the panoramic camera and VR tools are continuously evolving, shooting high quality video content is no longer a difficult thing to do.

Since omnidirectional video starts to become one of the main directions of next generation's rich media content, it is, of course, necessary to consider how to improve the user experience when a user is watching it through different channels. ODV can be viewed through different devices or setups such as personal computers, or mobile devices such as smartphones, HMD, CAVE or a dome (Benko & Wilson, 2010). One of the most valuable features of ODV is that it can provide immersive experience. And according to previous studies, it is most likely the HMD setup can have the best immersion effect (Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017).

On the other hand, ODV is a relatively new topic and still in lack of research literature. Until now, not many studies have looked into how basic ODV elements influence the user experience in detail. Hence, the thesis consists of an attempt to understand which height of the camera is an optimal choice for viewers to watch omnidirectional videos through HMD, and what range of height is acceptable for viewers, while they viewing the video in standing or sitting positions.

This thesis belongs to the field Human-Technology Interaction (HTI), which dedicated to promoting the use of technology by understanding how technology adapts to human needs. The main purpose of this thesis is to investigate the basic element of ODV, to understand user experiences and to give suggestions for future study.

This thesis focuses on three research questions related to 360-degree video:

- How does the height of the camera effect user experiences?
- What is the optimal camera height for videos which are viewed either sitting or standing?
- How does the user pose (sitting or standing) influence the user experience?

The research methods chose for this thesis study have all been selected based on these three research questions. In particular, a crucial concern of the thesis is the presence and immersive experience, which are subjective. In order to collect comprehensive user feedbacks, tasks with experience statements, questionnaires and semi-structured interviews were used in the user experiments.

The results of the approach described above include task performances logs, experience statements responses, comments, interview recordings and filled paper questionnaires. These data were summarized and analyzed with different qualitative and quantitative methods.

The remaining of the thesis has the following structure. Chapters 2, 3 and 4 provide more theoretical background related to the thesis, and further elaborate the research objectives and their meaning.

More specifically, chapter 2 presents an overview of the omnidirectional camera and 360-degree video. First presents the history of ordinary camera and the emergence of omnidirectional camera, then introduces the emergence of 360-degree video, after that some previous studies are reviewed, and finally challenges and guidelines regarding 360-degree video system design are discussed.

Chapter 3 is all about user experience. First, I distinguish between user experience and usability, then user experience evaluation methods are described, followed by studies related to the evaluation of 360-degree video system with an immersive environment or setups.

Chapter 4 introduces immersion and presence. First explains what *immersion* and *presence* are, then related studies regard 360-degree video system is introduced, finally explains why it matters.

Chapter 5 introduces the videos and the software used in the experiments. The video materials were provided by UTA's iMedia project team and the software used in the experiments was developed by a member of iMedia project as well. Detailed information about how videos were captured and how software works are presented.

Chapter 6 gives a detailed description of the conducted user experiments. The number of participants that were included in the experiment and participants' general information is given first. Then the experiments procedure and apparatus that were used in the experiments are introduced. Finally, how data was collected and analyzed is briefly discussed.

Chapter 7 reports the data analysis results and summary of the user study. Data sources include system logs (height selection), user experience statements, interviews and questionnaires. According to the characteristics of the data, different analysis methods have been applied.

In the end, chapter 8 and chapter 9 discuss the analyzed data results, provide suggestions for the future works and draw the conclusion.

2. Omnidirectional Camera and 360-degree Video

The 21st century is a new era full of with three-dimensional (3D) contents. Images, videos, audios, and interactive communication, all these elements together constitute the era of information. Among them, video image, mainly from the digital camera or webcam technology, plays a crucial role. Important advantages of video include intuitive, real, easy to spread, three-dimensional and practical nature. These make it becomes one of the most useful tools and means for people to record and communicate information to each other.

In recent years people have gradually started to use and love to use video shooting or camera as a basic life skill. People recording their daily activities as videos, posting them on the social media services, sharing these moments with family, friends or even strangers. These phenomena prove the importance of camera and video.

Regarding the devices that are used to shoot the videos, different types of camera have been used over the years. In this chapter, two types of cameras are discussed. The first one is the video camera, like its name tells, the camera is used for high quality video shooting. The second one, which will be introduced in detail later in this chapter, is the omnidirectional camera that is used for shooting the 360-degree videos. One of the existing omnidirectional cameras, Insta360 Pro, was used in the design and implementation phase of the thesis work.

This chapter starts with an overview of the camera development, more specifically, a brief history of the camera, and the emergence of omnidirectional video is introduced. Next, the evolution of the 360-degree video is given. Then, the purposes and objectives of ODV are explained to offer a better understanding of the emergence of the 360-degree video. Finally, some social impacts related to panoramic video and ODV are reviewed.

2.1 From Ordinary Camera to Omnidirectional Camera

From the invention of photography to the popularity of omnidirectional cameras today, the pursuit of capturing image or video accurately, easily and efficiently by people has never stopped. We can intuitively feel the change from film camera to digital camera, but all these still remain in the two-dimensional stage. With the emergence of omnidirectional camera, people are able to shoot a full spherical view, the limitation of shooting fragmented images and single direction videos has been broken. This section focuses on giving a comprehensive introduction to the camera and omnidirectional camera, starting with an overview of the history of the camera, then a general overview on omnidirectional camera is given.

2.1.1 The history of Camera

From ancient times to the present, the need for recording the wonderful moments of life has never stopped. In ancient times, people tried to record life with graphics and gradually evolved to reach the peak of the painting. To the 19th century, a French physicist, artist and a part time photographer (Wikipedia, 2017) Louis Daguerre invented the first practical silver version of the camera, more specifically, he created a photographic process which he named *daguerreotype*. On August 19, 1839, this photography was publicly announced on a joint rally held by both the French Academy of Sciences and the Academy of Fine Arts. Later French government renounced the patent for the invention and made it public (Kosinski & Kosinski, 1999). This day is usually taken as the beginning of the photography. At the end of the century the Kodak company in United States invented the world's first film machine which became the camera market standard. In 1975, Kodak created an operable electronic camera with CCD to lead the camera industry into the digital age (Wikipedia, Kodak DCS, 2018).

After that, the cameras have become more and more diverse and they can be classified from different angles. For instance, on a broad level, there are two types: professional camera, and general consumer camera. The former has a higher configurability, such as a lens with a better performance and a CCD with a larger size etc., thus it has a more prominent image quality and is easier to adapt to the environment. The latter is more suitable for home use, applications in the non-business areas of low image quality requirements, such as outdoor entertainment, family travel and so on. Since it has some advantages such as small size, light weight, portability, simple operation, and of course, cheaper price compared to the professional camera, it is a good choice for individual or family.

Figure 2.1 Cameras from old to new. Left: Very first camera; Right: Digital camera;

Since ordinary video camera has a limited field of view (FOV), it brings many restrictions while being applied in different applications (Nayar, 1997). For instance, in the navigation system, people need time to piece together the interior of a building from a limited-view video or the whole structure of a place such as a playground or touristic locations. Another example is the monitoring system with the traditional surveillance camera, it is difficult to monitor the entire area, as there is always a blind spot, which can be taken advantage by malicious people. This is why researchers have been working to find ways to broaden the FOV of the video camera, and the omnidirectional camera is exactly the result of the researcher's efforts.

2.1.2 Evolution of Omnidirectional Camera

As mentioned before, the omnidirectional camera is invented to extend the field of view. Nowadays many mainstream smartphones have the function of panoramic photo shooting, after turning on the camera function, users can simply use the smartphone to sweep a circle to get a panorama. However, the FOV of these panoramic photos is limited in one direction, either vertical or horizontal, this is why it cannot be considered as a "true panorama". A "true panorama" refers to a spherical image or video that covers a FOV of approximately 360˚ x 180˚, which refers to 360-degree horizontal and 180-degree vertical view angle.

With the explanation above it is possible to understand what is an omnidirectional camera and how it works, there are usually two ways to extend the limited field of view, one is using a single camera attached to a curved lens or curved mirrors, another one is to use multiple-camera setups in which cameras simultaneously record the environment and images are processed later (Bleumers, Van den Broeck, Lievens, & Pierson, 2012). Based on this concept, the omnidirectional camera is usually also known as panoramic camera, catadioptric camera, spherical camera, fish-eye camera or wide-angle camera (Scaramuzza, 2012). To some extent these synonyms reflect the main characteristics of the omnidirectional camera. As shown in Figure 2.2, different types of omnidirectional cameras exist.

Figure 2.2 Different types of omnidirectional cameras. Left: Catadioptric camera; Middle: Fisheye lens; Right: Spherical camera.

Despite the popularity in recent years, the history of omnidirectional camera is actually not that long. About 30 years ago, in the 1980s, scientists began to study how to acquire and use panoramic images. Back then, several studies related to robot navigation system, which aims to provide guidance for global navigation of a robot (Oh & Hall, 1987) were made. These studies all utilized a system called omnidirectional vision (navigation) system, two of the main components of the system are a CCD camera and a fisheye lens (or a lens with a very wide FOV) which provides a hemispherical field of view (Oh & Hall, 1987), this setup is intended to provide a largerFOV, reduce camera scanning and improve navigation efficiency of the robot. Later in 1990, for the same purpose, another research team proposed a new position sensor for mobile robot navigation system called COPIS (Conic Projection Image Sensor), in order to get an omnidirectional FOV, they used a glass pipe, a single camera with a perspective lens and a conic mirror (Yagi & Kawato, 1990), as shown in Figure 2.3, this setup looks unique but feels practical. In spite of immaturity, these setups have paved the way for the emergence of panoramic cameras.

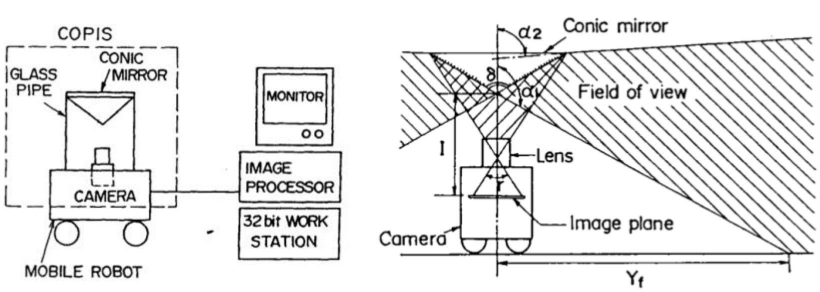
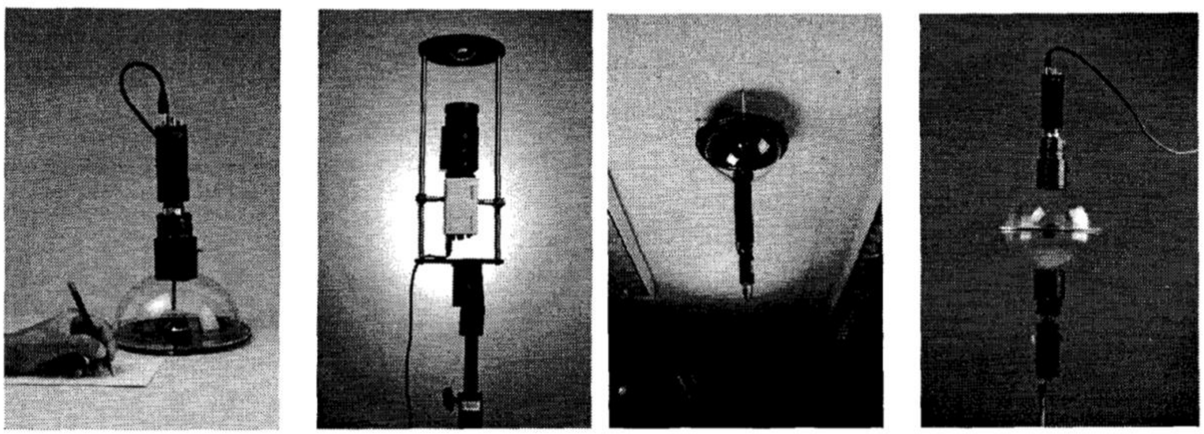


Figure 2.3 Configuration of COPIS, illustration of FOV of COPIS. Adapted from Yagi, Y., & Kawato, S. (1990)

However, one obvious problem with the setup described above is that it does not satisfy the single viewpoint constraint (Nayar, 1997). In the following years, scientists conducted further research on this subject and tried to propose more optimized solutions. Until 1997, Shree (1997) proposed a catadioptric omnidirectional camera concept based on the previous research, the main components of this camera are an orthographic lens and paraboloidal mirror (show in Figure 2.4), instead of the perspective lens in previous research. This results in a simplified calibration and computation of perspective images process. At the same time, the single viewpoint constraint is satisfied as well. After that, the omnidirectional camera had gradually been used in different fields, for instance, autonomous navigation, video conferencing and surveillance (Gaspar, Winters, & Santos-Victor, 2000; Delahoche, Pégard, Marhic, & Vasseur, 1997). In the same year, a FlyCam system for panoramic video imaging was presented by Foote and Kimber (2000). This system was also used in the study of extracting the region of interest (ROI) in the 360-degree video several years later (Sun, Foote, Kimber, & Manjunath, 2005), which can be useful for meeting indexing or face tracking. Only after this year, the size of omnidirectional cameras started to become miniaturized and their vertical field of view started to increase (Scaramuzza, 2012).



(a) (b) (c) (d)

Figure 2.4 Four omnidirectional cameras with paraboloidal mirrors. (a) For teleconferencing. (b) For navigation. (c) For surveillance. (d) Back-to-back configuration to sense the entire sphere of view. Adapted from Shree, K. N. (1997)

2.2 Emergence and Purposes of 360-degree Video

The 360-degree video, also known as omnidirectional video (ODV), spherical video or panoramic video, is the video that covers approximately a whole spherical field of view. As mentioned earlier, ODV is the video shot with the omnidirectional camera and covers a 360-degree field of view. It transforms the static panoramic images into dynamic video images, enables the user to freely watch the video at any direction he/she likes, and thus provides an immersive experience that will not be limited by spatial and temporal constraint. Due to new technologies such as omnidirectional cameras for recording and producing 360-degree videos, 360-degree videos have started to be widely known and accepted by people in recent years. Social media services such as Facebook and YouTube first brought the 360-degree video into the sight of the public, gradually, the 360-degree video starts to become widely accepted by various organizations and platforms and it has been utilized for different purposes.

The 360-degree videos can be viewed through different platforms or devices. Currently, many ODVs are designed to be viewed on ordinary 2D displays such as laptop or computer screen, or mobile devices screen (Petry & Huber, 2015). For example, ODV can be viewed via social media services such as YouTube or Facebook on a web browser or a smartphone. On the other hand, the technology introduced next can be better used for the design of interactive omnidirectional videos (iODV). Other three commonly used technology for watching ODV are the large screen, the tablet (Zoric, Barkhuus, Engström, & Önnevall, 2013) or spherical displays (Benko, Wilson, & Balakrishnan, 2008). The large screen can be used in different contexts, for instance, public activities (sports event or so) or cinema. Of course, there are more immersive visual setups such as Head-mounted displays (HMD) and cave automatic virtual environment (CAVE), both installations can provide a more immersive experience compared to the methods mentioned above. Although according to previous studies the HMD setup could provide the best feeling of immersion, both HMD and CAVE setups can enable navigating the natural space through head movement (Petry & Huber, 2015).

2.3 Previous Work

The reason why 360-degree video can be so popular is because of its unique advantages. This section will focus on some specific application processes and examples of 360-degree video in different fields.

Remote surveillance and monitoring is nowadays widely used in traffic control, home and office security, fire safety control and so on. In 1997, Peri and Nayar presented a real-time ODV exploration system called *OmniVideo* which they used for surveillance purpose. From the system, the user was able to freely choose the direction of an incoming ODV stream and, generate multiple perspectives and panoramic views from it. The parameters that were used to control the viewing direction, FOV or zooming could be adjusted via the interaction devices such as a mouse or a joystick (Peri & Nayar, 1997).

One year later, another research team in Japan proposed a new ODV surveillance and monitoring system which was a bit more complex (Onoe, Yokoya, Yamazawa, & Takemura, 1998). The system consisted of an ODV image sensor called *HyperOmni Vision* (it includes a hyperboloidal mirror and a standard camera), a video board, a workstation, wireless video communication setups and an HMD with a 3D head tracker, which means that besides choosing the viewing direction by clicking the mouse or by tracking moving objects in the video, the monitoring direction could be determined by the user's head movement as well. In the same year, the same scientific team proposed a telepresence concept called *instantour*, based on the concept they build a real-time virtual tour system (Onoe, Yamazawa, Takemura, & Yokoya, 1998), the hardware setup was basically the same as the setups mentioned in the last study. Since there was a wireless video communication setup, it was possible to transfer the ODV images from a remote place. The researchers tried to decrease the computation time (the time used to process ODV images) as much as possible, thus enabling the user to see ODV images of another location in real-time. It is worth to noticing that the system contained a multi-user mode which enabled multiple users to view different directions from a single point at the same time (Onoe, Yamazawa, Takemura, & Yokoya, 1998). With HMD, this system was intended to give users the feeling of being in another place.

In addition to surveillance and telepresence, all the systems described above can be applied in other situations, for instance, video conferencing, meeting monitoring (De la Torre, Vallespi, Rybski, Veloso, & Kanade, 2005), remote monitoring and autonomous navigation.

It is always challenging to provide an immersive experience for remote users. Regarding this problem, HMD seems like a good choice (Ochi, et al., 2014). To provide better immersive experience, Ochi and his partners (2014) designed an ODV streaming and viewing system. The system contained an ODV streaming part and a video viewing part. The former split the ODV into several small tiles for transmission, the latter is an HMD with an orientation sensor. Later in 2015, this system was put into practical use at the anniversary performance of a famous singer (Ochi, Kunita, Kameda, Kojima, & Iwaki, 2015), and this time the video viewing system was not limited to HMD, in addition to a web browser could be used. Another example is the use of ODV in the museum, which shows the ability of ODV to save memorable activities or events. In this case, two installations were presented (Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017), one is an HMD based rally simulator system, in this setup the user can sit in a real rally car and experience the rally race through HMD. Another is a CAVE-based road grader simulator system, the user is surrounded with a large screen and sits on a seat with haptic feedback, in this case, the user is able to control the steering wheel and the p (Benko & Wilson, 2010)edals, and see the corresponding changes in the big screen. This publication reports long-term studies related to ODV user experience, which will be described later in the User Experience section 3.3.

2.4 Challenges in Viewing 360-degree Video

Although the 360-degree video has very broad application prospects, and the emergence of modern technology such as HMD also provides a solid support for its development, its developing progress is inevitably accompanied with many challenges. As Kallioniemi et al. (2017) said in their research, because of the limited range of human FOV, the two main design challenges in ODV are "presentation of the content and interacting with it". After the appropriate expansion, this scope can be extended to the challenge of the video content itself, technical challenges such as shooting and viewing equipment, as well as interactive design challenges.

There are different kinds of challenges that may arise when trying to design systems that enable the user to interact with ODV content. In a study of ODV playback with HMD, researchers mentioned it is a challenge to design interactions for ODV playback control because the user cannot see his/her hands or actual environment (Pakkanen, et al., 2017). Zoric et al. (2013)summarized several points in their research. First it can be hard to balance the active and passive viewing when watching a 360-degree video. Sometimes the user may want to take the initiative to choose the viewing content, but sometimes the user will also worry about missing the important content and thus to give up the initiative and become passive. Then it is crucial to consider supporting orientation, as the user is free to explore the content of an ODV. It is worth considering how to orient the user after he/she is lost. How to give users a social experience while keeping their own viewing perspective is another challenge, which means users want to have their own viewing angle in the ODV, but still they want to share feelings and comments with other people. (Zoric, Barkhuus, Engström, & Önnevall, 2013)

In a first person omnidirectional video (FODV) system design project (Kasahara, Nagai, & Rekimoto, 2015), two technical challenges were mentioned. One is the weight and the comfort level of the HMD, and another is visual gaps in ODV images. At the same time, they also mentioned the video quality and the cyber sickness that may be caused by shaky video content, especially with devices which can provide more immersive experience, for example, HMD, CAVE or a large screen (Kasahara, Nagai, & Rekimoto, 2015; Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017). People may feel dizzy and nauseous when they find that the ODV content in HMD does not fit their physical situation or common sense, and thus their immersive experience is disrupted. In ODV system design and interaction, the lack of formulated design guidelines or specific evaluation criteria is another challenge that needs to be noticed.

3. User Experience of 360-degree Video

This chapter will start with the explanation of the differences between user experience and usability, then some user experience evaluation methods will be briefly introduced. In particular, those being used in this thesis work will be pointed out. At the end of the chapter, some studies and common issues related to ODV user experience will be covered.

3.1 User Experience and Usability

The term user experience was first widely recognized as being proposed and promoted by user experience designer Donald Norman (1998). User experience refers to a purely subjective feeling that is established during the process of using a product by a user. Its research focuses are the pleasure and value that the system brings, not the performance of the system. The exact definition, framework and basic elements of the user experience are still in development and evolving. However, the commonality of user experience for a predefined user group can still be recognized through well-designed user experience evaluation experiment, and thus applied in the future system design processes.

In recent years, the advances in mobile and graphics technology in computer technology have allowed the human-computer interaction (HCI) technology to penetrate almost all areas of human activity. This had led to a big shift: the evaluation indicators of the system extend from pure usability engineering to user experience enhancing. This has made user experience (user's subjective feelings, motivation, values, etc.) to be received considerable attention in the development of human-computer interaction technology. To some extent, it is comparable with traditional three major usability indicators which includes efficiency, effectiveness, and satisfaction (Frøkjær, Hertzum, & Hornbæk, 2000), or even more important. However, at the same time, it also poses a problem, the terms *user experience* (UX) and *usability* are usually getting confused and mixed up, while in fact, they are different concepts.

Although UX is getting more and more attention, there is still a lot of controversy over the scope and definition of UX (Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009). In ISO FDIS 9241-210 (International Organization for Standardization), UX is defined as: "*A person's perceptions and responses that result from the use and/or anticipated use of a product, system or service.*" On the other hand, the definition of usability has been defined or redefined several times throughout its development (Rusu, Rusu, Roncagliolo, Apablaza, & Rusu, 2015), but the most well-known and widely accepted one is also from ISO FDIS 9241-210: "*Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*" Through the definition of both, it is easy to see that the evaluation and measurement of usability need a specific product or system environment, and a context of use, while the UX does not.

There are different opinions regarding the relationship between usability and UX, which can be roughly summarized as follows (as shown in Figure 3.1): (1) UX is the refinement of satisfaction attribute in usability, (2) UX is the extension of the usability, (3) UX and usability are two separate concepts but affect each other. The theoretical background of these three opinions will be explained in more detail below.

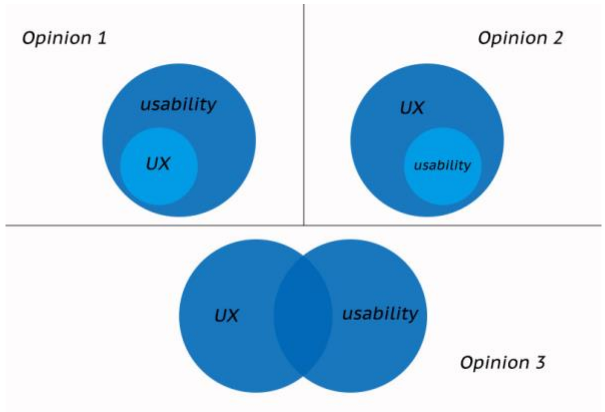


Figure 3.1 opinions regard user experience and usability

The first opinion considers UX as a purely subjective experience, a perception, a dilatation of satisfaction component in usability (Bevan, 2009a; Bevan, 2009b). In this case, UX is explained as a feeling of satisfaction, while using the product, if the user feels the product is efficient and effective to use, or user feels interested or enjoys the product, then it can be considered as a satisfying UX. To gain a comprehensive UX, both hedonic and pragmatic goals should be noticed. (Bevan, 2009a; Bevan, 2009b)

For the second opinion, usability is considered as one of the UX aspects (Väätäjä, Koponen, & Roto, 2009), or a core concept in UX (Lewis, 2014). In Väätäjä and her partner's research, a UX evaluation tool is presented, both hedonic and pragmatic goals are covered and qualified, and usability is included as one of the pragmatic qualities (Väätäjä, Koponen, & Roto, 2009). In Lewis's paper, UX is described as an extension of usability, so it also includes three indicators: effective, efficient, and satisfying (Lewis, 2014).

Finally, for the third opinion, Hassenzahl magnified the emphasis of each of UX and usability, thereby suggesting that these are two concepts with intersecting parts (Moczarny, De Villiers, & Van Biljon, 2012). For instance, UX exceeds usability in three aspects: (1) Holistic, usability focus more on performance of user's tasks while UX also focus on non-task parts such as personal perception; (2) Subjective, usability measures objective data includes task time or error rates, UX is more concerned with the user's subjective experience, the feelings about the system or product; and (3) Positive, usability tries to find out and fix the problem of the system, while UX also focuses on positive feelings and tries to amplify it (Petrie & Bevan, 2009). So, in contrast, usability surpasses the UX in these three areas: pragmatic, objective and performance (Moczarny, De Villiers, & Van Biljon, 2012).

For the third opinion, another study reached the same conclusion from a completely different perspective. In order to simplify the process of designing and evaluating technology, McNamara and Kirakowski (2006) proposed a three-element concept, as shown in Figure 3.2, which consists of *functionality*, *experience*, and *usability*. *Functionality* refers to the product itself, equipment, maintenance or so are within the scope. *Experience*, or user experience, refers to user's own feeling of using the product, is it comfortable or not, how is the perception etc. *Usability* is more about whether the product has reached the expectations of users and designers, the interactions, functions, reaction time and so on. In this study UX and usability are divided into two separate concepts, but they are not completely independent and to some extent affect each other. For instance, a bad usability is bound to result in a poor user experience, but a good user experience does not necessarily lead to good usability. (McNamara & Kirakowski, 2006)

Figure 3.2 Three components of technology design and evaluation. Adapted from McNamara & Kirakowski (2006)

It is obvious that usability is more task-oriented and requires a specific context of use (Petrie & Bevan, 2009). However, UX is freer about context of use, it is more practical for emerging technologies with diversified product orientations. This explains why UX evaluation methods will be used in this thesis work, because there is necessarily no set of goals for participants to achieve in the study, and no specific tasks need to be completed. The study focuses more on the UX and feelings while using the technology.

3.2 Evaluation Methods

This section starts with usability evaluation methods and leads to UX evaluation methods. I briefly explain the difference between usability evaluation methods and UX evaluation methods and present some UX evaluation methods that are used in this thesis study.

3.2.1 Usability Evaluation Methods

Until UX becomes an independent concept, system or product design and evaluation tend to focus on usability. In terms of the definition of usability mentioned above, this means that the evaluation of usability includes the evaluation of effectiveness, efficiency, and satisfaction. In addition, several other aspects proposed by some researchers such as flexibility (the degree to which the system adapts to change), learnability (the time user spends learning systems functions), memorability (the time it takes for the user to return to the system after a period of time), safety (whether the system can ensure user safety) (Petrie & Bevan, 2009) and error rate (error rate when user using the system) (Rebelo, Noriega, Duarte, & Soares, 2012) may be also included in evaluations based on the situation.

In order to evaluate these aspects, researchers need to plan ahead to determine the most effective and efficient evaluation method. There are different ways to categorize usability evaluation methods, Fernandez et al. (2011) mentioned one way of classification which is easier to understand.

They proposed that the usability evaluation methods can be broadly divided into two broad categories: empirical methods and inspection methods (Fernandez, Insfran, & Abrahão, 2011). The former focuses on gathering data from users, for instance, observe users while they interact with the product or product prototype, define a series of tasks for user to complete, during which usage data includes task completion time or frequency of errors will be collected (Fernandez, Insfran, & Abrahão, 2011; Rebelo, Noriega, Duarte, & Soares, 2012). The latter, on the other hand, focuses on the product itself and does not require user participation. Such methods include product interface review based on a series of guidelines such as usability heuristics (Nielsen, 1995), or automated usability checking tools used after the product or product prototype is fully implemented (Petrie & Bevan, 2009).

3.2.2 User Experience Evaluation Methods

Although the two are easily tangled from the standpoint of usability and UX definition, the UX evaluation methods can be quite different from usability evaluation methods. Since UX concept includes the understanding and optimizing of user expectations, preferences, feelings (emotions and enjoyment), and needs (Bevan, 2009; Rebelo, Noriega, Duarte, & Soares, 2012; Obrist, Roto, & Väänänen-Vainio-Mattila, 2009), at the same time it is also context-dependent and dynamic (Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009). This means that even with the same product or system, the UX can be very different in different environments. Therefore, for UX researchers, how to establish useful evaluation methods for subjective measurements is a topic that has always been focused.

By now, many effective methods for UX evaluation have been proposed and widely used. In a user experience evaluation methods overview, totally 123 UX evaluation methods were collected from different sources, including workshops, SIGCHI (Special Interest Group on Computer-Human Interaction) conferences and some other sources. Some duplicated methods or methods that are not considered to be in the UX scope were excluded, so in the end totally 96 UX methods were screened out and listed on the web at http://www.allaboutux.org/all-methods (Vermeeren, et al., 2010). In SIG on CHI'09, participants were asked to answer the question: "*User Experience Evaluation – Do you know which Method to use?*", and finally 36 methods were collected, the categorized result can be seen in Figure 3.3 (Bevan, 2009; Obrist, Roto, & Väänänen-Vainio-Mattila, 2009; Vermeeren, et al., 2010). Among which, some common and easy-to-use methods such as retrospective interview, questionnaires (with rating scales) or survey questions (including Emocards or Emofaces) are included. In this thesis study, various means including questionnaires, semi-structured interview and rating scale questions (including Emofaces) were combined together to evaluate user experience during the experiments.

Evaluation context
Lab tests
Lab study with mind maps
Paper prototyping
Field tests
Product / Tool Comparison
Competitive evaluation of prototypes in the wild
Field observation
Long term pilot study
Longitudinal comparison
Contextual Inquiry
Observation/Post Interview
Activity Experience Sampling
Longitudinal Evaluation
Ethnography
Field observations
Longitudinal Studies
Evaluation of groups
Evaluating collaborative user experiences,
Instrumented product
TRUE Tracking Realtime User Experience
Domain specific
Nintendi Wii
Children
OPOS - Outdoor Play Observation Scheme
This-or-that
Approaches
Evaluating UX jointly with usability

Evaluation data
User opinion/interview
Lab study with mind maps
Quick and dirty evaluation
Audio narrative
Retrospective interview
Contextual Inquiry
Focus groups evaluation
Observation \ Post Interview
Activity Experience Sampling
Sensual Evaluation Instrument
Contextual Laddering
Interview
ESM
User questionnaire
Survey Questions - Emocards
Experience sampling triggered by events, SAM
Magnitude Estimation
TRUE Tracking Realtime User Experience
Questionnaire (e.g. AttrakDiff)
Human responses
PURE - preverbal user reaction evaluation
Psycho-physiological measurements
Expert evaluation
Expert evaluation
Heuristic matrix
Perspective-Based Inspection

Figure 3.3 UX evaluation methods from CHI'09. Adapted from Bevan, N. (Bevan, 2009)

3.3 User Experience Evaluation for 360-degree Video System

As mentioned in the second chapter, 360-degree video, thanks to its good user experience and interaction capabilities, has been widely used in many fields, such as industrial design, entertainment, education and so on. With the increasing demands for good user experience, how to evaluate and provide better user experience of 360-degree video in various applications now becomes the next hot topic.

In a recent study, a research team demonstrated the result of their long-term user experience study with two 360-degree video based museum installations (Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017). One is a rally simulator based on HMD, there is not much interaction in this setting however the user can sit in a real rally car while using the device. Another is a road grader simulator with three projectors, the user can sit in a physical replica of a grader interior with three projectors images in front of him/her. This setup is more interactive since the system can respond to user's behavior (stepping on the pedals or turning the steering wheels) and the seat of the simulator can provide haptic feedback when detecting related trigger events as well. To gather the subjective feedback, which includes the effects of different settings on the user experience, immersive experience, interactive experience and the influence of modalities, from the two installations, two similar questionnaires (one for each setting) were used during the feedback gathering period. Each questionnaire has around ten questions and was designed to have five rating scales, from *Totally disagree* to *Totally agree*. Considering that most of the target users are children, the questionnaire was in paper form. The results of the user study show that overall, it is fun to use both of the simulators. More detailed feedback indicates that although HMD installation can provide better immersive experience compared to the other, it can cause motion sickness more easily as well. (Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017)

In another study which aims to figure out the differences in the user experience and immersion between HMD and CAVE applications, a special user experience evaluation method called

SUXES was used (Kallioniemi, et al., 2017; Turunen, et al., 2009). SUXES was designed to gather subjective user experience feedback from multimodal applications. The unique part of this evaluation method is that the user needs to fill in an expectation questionnaire before using the application and an experience questionnaire after actually using the application. The questionnaires contain several Likert scale questions. The method was used in the HMD and CAVE experiments and the results were pretty positive, it seems that user's actual experience surpasses the expectations in almost all statements about the application especially with the HMD setting. As for the immersive experience, the feedback shows that HMD can provide a higher degree of immersion and users tend to spend less time in completing the task with the HMD. (Kallioniemi, et al., 2017)

4. Immersion and Presence

With the advent of concepts and technologies such as virtual reality, virtual environments, 360-degree video or head-mounted displays, the terms *immersion* and *presence* have also emerged in a growing number of news and research papers. Immersive experience and sense of presence are usually considered as important features while talking about virtual reality applications, and also, 360-degree video applications. However, these two terms are normally quite vague and hard for people to tell the difference. This may not influence the user, but for a designer and a researcher, no understanding the difference may make the design of better accepted and easy to use product or system more challenging. Knowing the difference between the two terms can provide a better understanding of how to design for or evaluate them.

Since immersion and presence starts to be noticed and applied in different contexts, various methods or techniques have been developed to evaluate immersion and presence. Common approaches include subjective measures such as questionnaires and interviews, and objective measures includes behavioral and physiological responses.

The current chapter first gives various opinions regarding the definition of immersion and presence, then based on previous studies, several examples about the effects and applications of immersive experience and presence, and how to measure or evaluate immersion and presence are given. The examples involve both virtual reality applications and omnidirectional video applications since they bring the similar experience to some extent, especially when using technologies such as the head-mounted display or dome environment. Finally, a summary is given.

4.1 Definitions of Immersion and Presence

Because the term immersion is often used as a synonym to presence and results in confusion (McMahan, 2003), researchers attempt to differentiate the two terms. Among multiple opinions, there are two prominent views that distinguish the immersion and presence, and the biggest difference between these two concepts is the definition of immersion. One defined immersion as an attribute or function of a system, in other words, a technology, and another defined it as a subjective or psychological experience (Nilsson, Nordahl, & Serafin, 2016).

The idea that immersion is a technological property is provided by Slater and Wilbur (Slater & Wilbur, 1997), they defined *immersion* as:

- *"A description of a technology, and describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant."*

And *presence* as:

- *"A state of consciousness, the (psychological) sense of being in the virtual environment."*

On the other hand, the opinion that immersion is a perceptual feeling is proposed by Witmer and Singer (1998). According to their research (Slater, 1999; Witmer & Singer, 1998; Mütterlein & Hess, 2017), *immersion* is defined as:

- *"A psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences."*

And *presence* is defined as:

- *"The subjective experience of being in one place or environment, even when one is physically situated in another."*

4.1.1 Presence

Basically, in these two studies, the definitions towards presence are similar, both refer to a subjective sense of 'being there', where 'there' can indicate either a virtual environment that is generated by computer and does not exist at all, or a 'real' environment that is different from current physical environment (for example, a distant environment viewed through an omnidirectional camera). However, the two research teams have very different approaches to expand the content of the *presence*.

According to Slater and Wilbur, presence has a strong correlation with immersion, based on their understanding. Presence can be seen as an extension (or extended response) of immersion, and three different aspects of presence were mentioned in their study (Slater & Wilbur, 1997): (1) Presence contains both subjective and objective parts, the subjective part refers to how strong the feeling of 'being there' in a virtual environment. While the objective part indicates how similar is the user's pattern of behavior in the virtual environment compared to similar situation in reality; (2) The higher the presence level, the more influence the virtual world has on the user compared to the real world; (3) After visiting the virtual environment, to what extent the users think the virtual environment as 'an actual place' rather than just some generated images. It is noticeable that Slater and Wilbur connected presence and immersion with an equation: presence (here it means the response or the measurement result of immersion) is placed on the left-hand side and the immersion (here it means system characteristics) is placed on the right-hand side. This brings big challenge of measuring presence itself, since Slater and Wilbur believe the differences between individuals will influence the variables in the equation (Slater & Wilbur, 1997), even in the exact same experiment setup, the results can be totally different between two participants due to their personalities.

On the other hand, Witmer and Singer interpreted presence from a very different perspective. In their study, several factors that may influence the presence were mentioned (Slater, 1999; Witmer & Singer, 1998): (1) Control factors, which contain degree of control, immediacy of control, anticipation, mode of control and physical environment modifiability; (2) Sensory factors, which contain sensory modality, environmental richness, multimodal presentation, consistency of multimodal information, degree of movement perception and active search; (3) Distraction factors, which contain isolation, selective attention and interface awareness; (4) Realism factors, which contain scene realism, consistency of information with the objective world, meaningfulness of experience and separation anxiety/disorientation. All these factors are related to subjective feelings, and based on the defined factors, Witmer and Singer (1998) designed a Presence Questionnaire (PQ) to measure the presence experienced in a virtual environment and how potential contributing factors influence the participant's overall experience. It is noteworthy that Witmer and Singer also believe that the differences between individuals and the characteristics of the virtual environment will influence the degree of presence, so they developed the Immersive Tendencies Questionnaire (ITQ) to evaluate the capabilities of the participant (Witmer & Singer, 1998).

4.1.2 Immersion

As mentioned earlier, the most prominent difference between these two concepts is the definition of *immersion*. Slater and Wilbur defined it from an objective point of view, saying that the technology or the actual virtual environment system is the key point of deciding to what extent the immersive experience can be offered (Slater, 1999), in other words, *immersion* describes to what extent the system can offer participant an inclusive, extensive, surrounding and vivid sense of reality (Slater & Wilbur, 1997). Here are some examples that can help to better understand this opinion. For instance, consider two given systems, on the premise that all other conditions are the same, if one has a better display resolution than the other, then the former is considered more immersive compared to the latter. Again, with two systems, if one accommodates more sensory modalities (such as light, sound, temperature, pressure) than the other, then the first is considered more immersive compared to the second.

Witmer and Singer, however, considered *immersion* as a subjective concept, they also mentioned several factors that would influence immersion in virtual environment in their study (Witmer & Singer, 1998; Nilsson, Nordahl, & Serafin, 2016): (1) Isolation from the physical world. Given a system, if it can effectively separate the users from the real word environment, it tends to increase the degree of immersion; (2) Perception of self-inclusion in the virtual environment. If the users feel like they are out of the virtual environment and looking at some generated images, they tend to lose the feeling of immersion; (3) Natural modes of interaction and control. When the users are able to naturally interact with the virtual environment, the degree of immersion can be increased; (4) Perception of self-movement. If the users feel they are actually moving in the virtual environment or interacting with objects inside it, their immersion level is likely to be increased.

Overall it is easy to see that Witmer and Singer's understanding of *immersion* is actually part of Slater and Wilbur's understanding of *presence*. In Witmer and Singer's paper (Witmer & Singer, 1998), they expressed their disagreement with Slater and Wilbur's view of immersion and, emphasized that immersion is a personal experience, just like presence. As a response to this view, and also to avoid unnecessary confusion in future research, Slater decided to use term *system immersion* to indicate his own understanding of immersion, and *immersion response* to represent Witmer and Singer's notion of immersion.

Despite the differences in these two opinions, they both include one crucial point: the necessity of mapping participant’s virtual experience with physical experience. Slater and Wilbur’s system immersion targets the degree of realism of user’s surrounding: the more realistic the user consider the environment to be, the stronger the sense of immersion. Witmer and Singer emphasize user’s self-consciousness and user's interaction with the virtual world: the more user's behavior patterns correspond to physical behavior patterns, the stronger the sense of immersion.

4.2 Previous Work

Immersion and *presence* are important features for ODV and virtual reality applications. This section first introduces some immersive ODV applications, then presents some methods used for immersion and presence measurement. Lastly, some research related to symptoms that easily appear in immersive installations are mentioned.

4.2.1 Applications of Virtual Environments and ODV System

Several years ago, a Brussels theatre company CREW supported a performance called C.A.P.E (Cave Automatic Personal Environment) with immersive technologies (Decock, Van Looy, Bleumers, & Bekaert, 2014). By using an ODV system, they tried to lead users into another world and experience the place that they have never been to before. The users were wearing an HMD and were accompanied with a CREW buddy, who helped the user and assisted by providing some tactile feedback. In order to reduce incoherence in the narrative process, or in other words, the cuts between the scenes, a CREW buddy shifted the user's attention by touching or pushing him/her at some specific moments (Decock, Van Looy, Bleumers, & Bekaert, 2014). To increase the sense of presence, various factors that are considered to affect user experiences such as camera techniques, interactivity, image and audio qualities were all optimized in the performance. Users' feedback was gathered through pre-questionnaire and post-questionnaire. The questions were related to different factors of presence. The analyzed data showed that users enjoyed the performance and had an immersive experience. However, the level of the enjoyment was related to whether they could identify themselves as the protagonist (Decock, Van Looy, Bleumers, & Bekaert, 2014).

In another piece of research, an ODV-based immersive installation was presented. The immersive experience was provided by an interactive dome, providing a 360-degree field-of-view, which was projected by a wide-angle lens projector (Benko & Wilson, 2010). The dome allows multiple users to enter at the same time and use hand gestures to interact with it. Different hand gestures including pinch, two-hand-circle and hand clasp were used to increase the immersive experience. Furthermore, the researchers proposed that the laser pointer or the shadow of the user can be applied as interaction tools (Benko & Wilson, 2010).

4.2.2 Measurement of Immersion and Presence

Various methods can be used to evaluate immersion and presence. Schuemie et al. (2001) did a survey about presence in virtual reality and concluded some measurement methods for presence. For subjective measurement, several typical questionnaires were mentioned, including Wilbur and Singer's ITQ and PQ, Schubert et al.'s IPQ (Igroup Presence Questionnaire) and ITC-SOPI (ITC Sense of Presence Inventory). Other subjective measurements include conducting small focus group discussions after experiencing immersive environments or the use of devices to continuously measure the level of presence (Schuemie, Van Der Straaten, Krijn, & Van Der Mast, 2001). Presence can be measured objectively as well, for instance, by observing user's subconscious behavior or the change of user’s physiological status (Schuemie, Van Der Straaten, Krijn, & Van Der Mast, 2001).

Combining questionnaires with other evaluation methods seems to be the most widely used approach for immersion and presence evaluation. In a study that aimed to measure the immersive experience in games, questionnaires were combined with tasks and eye-tracking devices (Jennett, et al., 2008). Both quantitative and qualitative data were analyzed, and the results indicate that both positive and negative feelings of immersion exist (Jennett, et al., 2008). Another research compared the user experience and the level of immersion between CAVE and HMD. The applications used in both installations were based on interactive ODV (Kallioniemi, et al., 2017). Evaluation method SUXES (Turunen, et al., 2009) was applied, the data from pre-questionnaire (expectation questionnaire), task (task completion time) and post-questionnaire (experience questionnaire) were collected and summarized in the study.

4.2.3 Issues that Break Immersion and Presence

Despite the fact that ODV can bring presence and the immersive experience, it often comes with the problem of cyber sickness. Researcher Palmisano et al. (2017) investigated the cyber sickness caused by mismatches between visual effects and actual head movements under HMD installation. They designed three different conditions in the experiment to validate the hypothesis: whether the visual compensation is ‘compensated’, ‘uncompensated’ or ‘inversely compensated’ for actual head movements (Palmisano, Mursic, & Kim, 2017). The results confirm the hypothesis, the cause of cyber sickness is connected to the mismatches between visual and physical head movements. In a research that studied the user experiences of ODV applications in the museum context (Hakulinen, Keskinen, Mäkelä, Saarinen, & Turunen, 2017), the camera movements were stabilized as much as possible to generate stable videos in order to reduce the cyber sickness. The positive result indicates that the occurrence of cyber sickness can be decreased by matching user’s visual aspects with physical head movements as much as possible.

5. Evaluation Materials and Software

This thesis investigates the influence of camera height of ODV on user experience when viewing in sitting or standing position using and HMD. To evaluate this, a set of videos were shot, and a software was developed.

This chapter describes the videos and the software used in the evaluation. The videos were captured and provided by the members from the iMedia project. The author participated in and early test video capture session. More information about the project can be found in later sections. At the beginning of the chapter, the process of how the videos were recorded and the content of the videos is briefly introduced. The software was developed by a member of the iMedia project as well, the author and the developer agreed on what should be included in the software and how it should work. A comprehensive description of the software is given at the end of the chapter.

5.1 Videos Used in Experiments

The video materials that were used during user experiments were created in iMedia project (Immersive Media project, TAUCHI, University of Tampere). The project aims to investigate how different factors of ODV influence user experiences. The factors include infographics in ODV, actions of persons in ODV, viewing position (sitting and standing), camera height and so on. Since the object of my thesis work is to study the effect of camera height of ODV on user experiences, videos that were shot for iMedia project could be used in this thesis study as well. To reduce the extra effort of shooting videos, my supervisor and one of the members of iMedia project, Jaakko Hakulinen, gave me the permission to use the same videos in my thesis study. One point that needs to be mentioned is that in iMedia project, user experiments focus more on objective data, behavioral measure collected in iMedia evaluations include head movements and eye tracking data are collected during the experiments. My thesis study, on the other hand, involves more subjective data, sense of presence and immersion were gathered through questionnaires and small interviews. This way we can produce various views for the same topic from different angles.

In order to eliminate the variables that may have an impact on the UX, the varying elements in the videos were minimized. The videos did not contain any sounds. As shown in Figure 5.1 and Figure 5.2, the videos were shot in a closed room, there are four actors in the video, standing in the four corners. Behind each actor there is a table with a piece of paper on it. The omnidirectional camera was set in the center of the room. The only varying element in the videos is the height of the camera. As Figure 5.2 shows, there are in total 19 camera heights, 4 for presence and immersion evaluations and 16 for height selection. Each video corresponds to one camera height, camera height was adjusted every time after a video was shot.

Figure 5.1 Actual setups for video shooting (4 actors standing in the four corners)

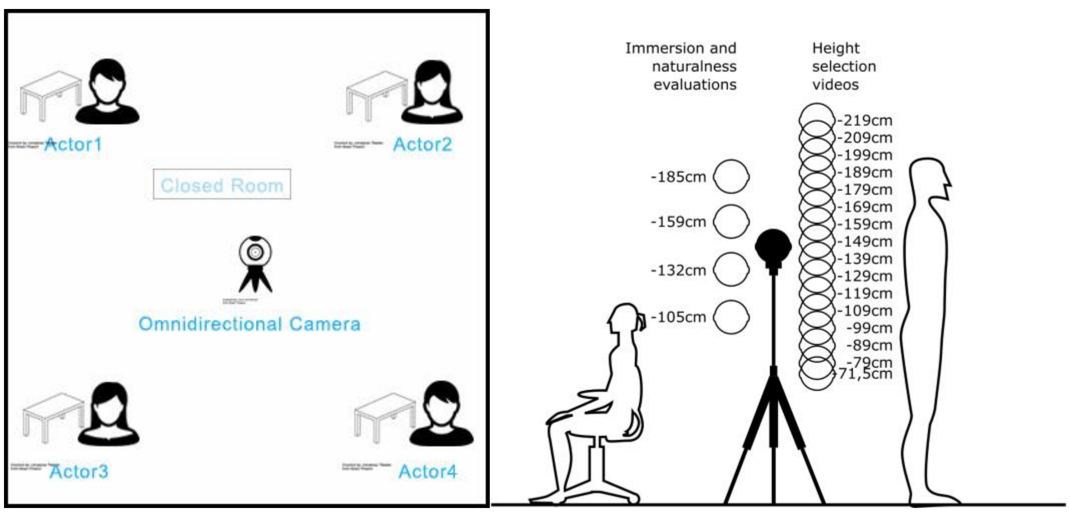


Figure 5.2 Left: Setups for video shooting; Right: Camera height for different videos

For the purpose of making video content less boring and providing the user with visual references, in the video, the actors perform two rounds of movements, each round last for around 20 seconds. The movement path is shown in Figure 5.3. To make sure that the user explores the ODV from all directions, the content of the videos is designed to enable a small task. After one round of movement, each of the actors picks up the paper with a mark on it from the table. As demonstrated in Figure 5.4, there is one paper with a cross ‘x’ on it and the others have the circle ‘o’, this is intended to let the user find the ‘x’ one in the experiments.

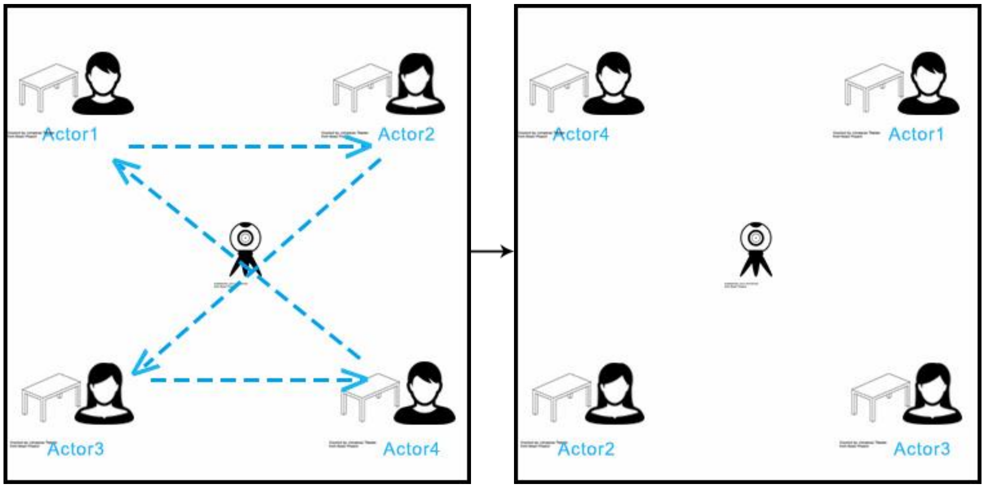


Figure 5.3 Left: Moving path of the actors; Right: Positions after the first movement

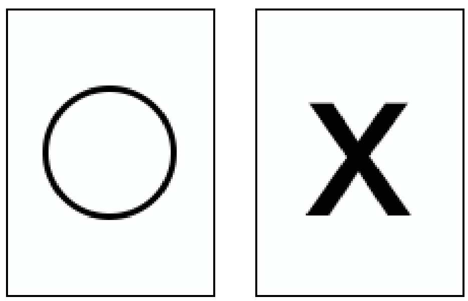


Figure 5.4 Paper with ‘o’ and ‘x’ on it

All the videos were shot by Insta360 Pro Camera, the camera was selected because it is able to record maximum at 7680x3840 high-resolution videos, easy to control and review the videos from the phone, or edit videos on personal computer. The original videos for evaluation purpose were recorded in 6400*6400 (6K) resolution, the videos were processed and finally 4096*2048 (4K) resolution videos were used in the user experiments. The camera is mounted on an adjustable tripod stand. After each shooting, the height of the tripod was adjusted to reach the corresponding camera height.

For immersion and presence evaluation, the four heights were selected because they were considered as the appropriate heights both for female and male, and cover sitting and standing heights of most people. As the average adult human height in Finland was around 180cm for male and 165cm for female according to the measurement between 2010 to 2011 (Wikipedia, 2018). It is noticeable that the heights mentioned here were measured based on the camera lens height, not the top of the camera.

5.2 Software Used in Experiments

The software used during the experiment process was developed by Kimmo Ronkainen, one of the members of iMedia project. The software was developed using Unity3D, a powerful game development platform, suitable for developing 3D software.

The software mainly contains two parts and the contents of the two parts are basically similar. The two parts correspond to the two viewing conditions of standing and sitting respectively. The content flow of the software can be seen in Figure 5.5. For each part, first, a transitional video was shown, the transitional video was used for users to familiar with the headsets and understand how it works. For this evaluation the transitional video contains a short video clip that was shot in a park in Tampere, Finland, note that the transitional video contains a bit sound. Next four height video sections with different camera heights were viewed. Each video section included the height video and Likert scale and Emofaces evaluations of the resulting user experience. After that the transitional video was shown again. Then a special video that contains 16 different heights was viewed. Then as the end of the first part and the beginning of the second part, the transitional video is added again. At last, the similar procedure was repeated in the second part.

There was a file named *config.txt* in the software file system. The file can be edited to determine the conditions in the evaluation experiments. The output of the software was recorded in a file called *log.txt*. More about the conditions of the evaluation can be found in Section 6.1.

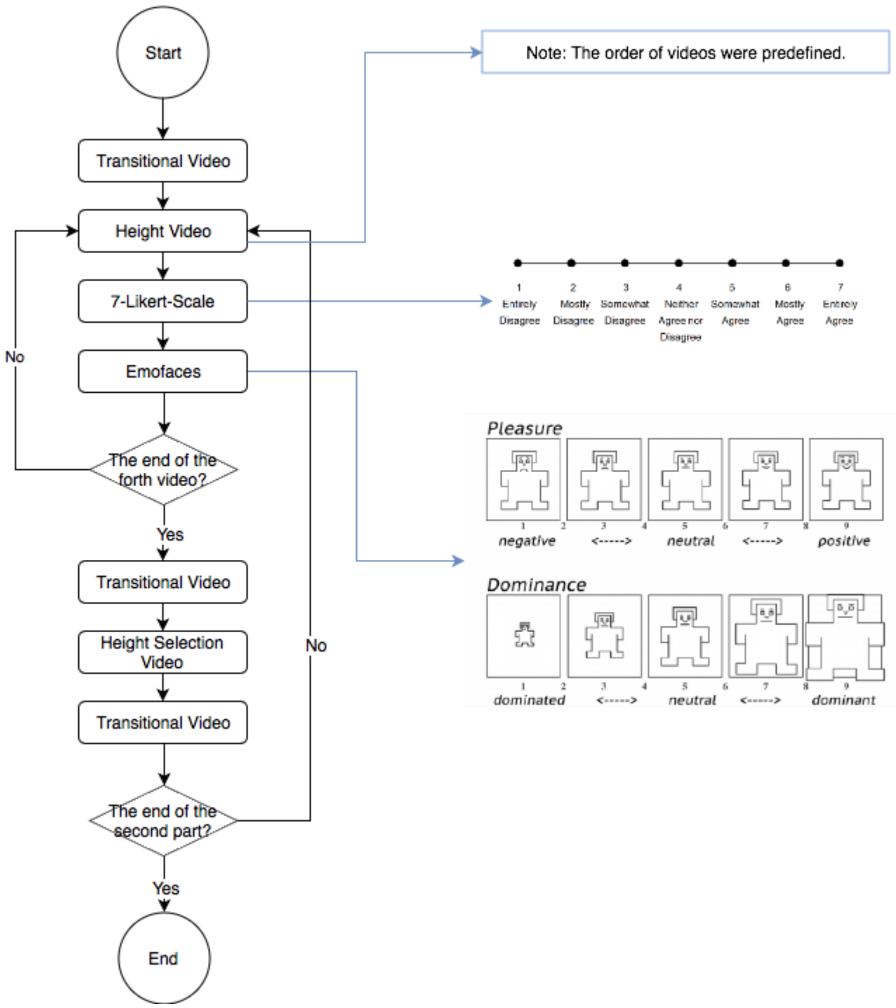


Figure 5.5 The content flow of the software

As shown in Figure 5.5, the 7-point Likert scale and the Emofaces are pictures that were displayed to the participant. To make sure that user can see pictures from a horizontal viewing angle while using the ODV software, the pictures were placed horizontally in three different directions. As depicted in Figure 5.6, the blue dots represent the central point of pictures, the central points of three pictures can form a triangle.

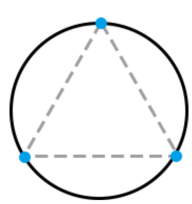


Figure 5.6 Horizontal view sketch of ODV software

6. User Study

A within-participant arrangement was chosen for the experiment. Each participant was exposed to the two viewing positions: standing and sitting. The participants were asked to use the software, interviewed during and after the experiment and were asked to fill in questionnaires. To ensure comprehensive feedback, interviews were recorded, and usage information of the software was logged. The audio recordings were transcribed into texts to extract useful comments and subjective feedbacks.

This chapter first describes the participants and background questionnaire, then introduces the apparatus and the procedure of the evaluation, and the presence questionnaire and immersive intendancy questionnaire that used in the experiments are presented. At last, data collection and analysis are briefly described.

6.1 Participants

A total of 16 participants took part in the experiments. There were 9 females and 7 males with the average age of 25.4 (SD = 3.63). English was used as the language of instruction for all the participants regardless of their native language. The participants were recruited from the University of Tampere and Tampere University of Technology campuses and were compensated with a movie ticket for their participation.

Table 1 summarizes the participants' demographics. 8 participants were wearing eyeglasses during the experiments, and there was 1 participant who was not wearing eyeglasses but wearing the contact lenses. 10 participants had prior experience using VR headset, 3 of them had used VR headset more than 5 times a year, 2 mainly for project works and 1 mainly for watching movies. The rest of the participants had only used VR headset 1 or 2 times a year. The experiences came from visiting museums, playing games, participating in user studies or demo presentations. 12 participants had watched 360-degree videos before, 2 had watched over 5 times a year and were mainly watching for entertainment and project work, 4 had watched 3 to 5 times a year, for entertainment, project work, and research study, the remaining 6 participants had only watched 1 or two times, for entertainment or education purpose. The main viewing channels for these panoramic videos included mobile devices (4 participants mentioned), personal computers (2 participants mentioned) and VR headset (7 participants mentioned). There were 5 participants who had experienced the feeling of motion sickness before, 2 were due to using VR headsets, the other 2 were felt while watching game videos or rotating scenes, the last one had felt dizzy while watching 3D movie in the movie theatre.

Participant Number	Age	Gender	Height	Wearing Eyeglasses?	VR Headset Experience	ODV Experience	Motion Sickness Experience
15	24	Male	178cm	Yes	No	no	no
11	37	Female	160.02cm	Yes	Yes	yes	yes
14	25	Female	164.59cm	No	Yes	yes	yes
13	20	Female	163cm	No	No	Yes	yes
9	23	Female	175cm	No (Contact lenses)	Yes	Yes	yes

12	26	Female	165cm	No	Yes	Yes	no
10	24	Male	183cm	Yes	Yes	No	no
16	27	Male	168cm	Yes	Yes	No	No
3	23	Female	162cm	Yes	No	Yes	yes
5	26	Female	160cm	No	Yes	Yes	No
4	25	Male	178cm	No	No	Yes	No
2	26	Female	154cm	Yes	Yes	No	No
8	28	Male	182cm	Yes	Yes	Yes	No
7	26	Male	186cm	No	Yes	Yes	No
1	24	Female	155.45cm	Yes	No	Yes	No
6	23	Male	183cm	Yes	No	Yes	No

Table 6.1 Background information of participants

As mentioned before, in total 19 videos were recorded. As Figure 5.1 shows, only 4 of the videos were used to evaluate the presence and immersive experience, and 16 of the videos were used in the height selection part (see section 6.2). To counterbalance the conditions in the evaluation experiments, the conditions were arranged using Latin Square, the sitting and standing conditions were included in this ordering to minimize switches between them. The resulting ordering is listed in Figure 6.1.

A = height 1 video
B = height 2 video
C = height 3 video
D = height 4 video

condition 1: sitting
condition 2: standing

Participant 1: sitting: A B D C standing: B C A D
Participant 2: sitting: B C A D standing: C D B A
Participant 3: sitting: C D B A standing: D A C B
Participant 4: sitting: D A C B standing: A B D C
Participant 5: standing: A B D C sitting: B C A D
Participant 6: standing: B C A D sitting: C D B A
Participant 7: standing: C D B A sitting: D A C B
Participant 8: standing: D A C B sitting: A B D C

Figure 6.1 Counter balanced conditions of the evaluation experiments

6.2 Procedure and Apparatus

The experiments took place in SimLab, a laboratory located at the University of Tampere. Before each participant arrived in the laboratory, the author checked the software and the instruments, made sure they were ready for the experiments, then edited the config.txt file in the software file system by entering the number of the participant, so that the software could display the videos in the correct order.

When a participant arrived the laboratory, he/she was greeted and asked to have a seat. The participant was then informed of the goals of the experiment and procedure and asked to sign the informed consent form. After signing the consent form the participant was asked to fill in a background questionnaire. Appendix D shows the content of the background questionnaire.

After completing the background questionnaire, the more detailed information about the procedure, including tasks, was given. The participant was familiarized with the apparatus, including HTC VIVE headset and controllers, as shown in Figure 6.3. There is a trackpad and a trigger on the controller, which were used in completing the tasks.

Figure 6.3 HTC VIVE. Left: headset; Middle and right: controller

The participant was then invited to sit in a rotating chair or stand in the center of the room, depending on the conditions assigned, and to put the headset on. The author assisted the participant to adjust and calibrate the headset. After all preparations were completed, the transitional video (i.e., practice video) was put on. The participant was notified to freely explore the video and when he/she felt he/she was ready for the actual tasks, a sequence of videos with different conditions was put on. Figure 6.4 shows the experiment environment.

Figure 6.4 A participant exploring the ODV.

The tasks were to find the paper with the icon ‘x’ in every single video. The participant was told to say something like ‘ok I found it’ and press the trigger on the controller when finding the correct icon in the video. Each time after viewing the video and completing the task, 7 Likert scale statements were verbally asked one by one from the participants to evaluate the experience of the previous video. After all the 4 videos had been viewed and all the statements for videos had been asked, the participant was told to prepare for the height selection part.

The height selection video actually consists of 16 videos (height information showed in Figure 5.1). The participant could use the controller’s trackpad to adjust the camera height. Pressing the upper part of the trackpad increased the height and pressing the lower part of the trackpad decrease the height. The participant was asked to pick out the optimal height video for four circumstances based on their feeling: (1) The most natural one; (2) The most enjoyable (comfortable) one; (3) The lowest acceptable one; (4) The highest acceptable one. After the participant had selected the optimal videos for these four situations, the participant was asked to take off the headset and was interviewed with two questions: (1) *How is your feeling now?* (2) *Did the environment make you feel cyber sickness (or nausea， dizzy) at some point? If yes, please specify.* Further questions could be asked based on participants’ performance. Then he/she was told to take a small rest before the second section started.

In the second phase of the experiments, the participants who had started the experiment sitting down, were then invited to stand up and stand in the center of the laboratory throughout the experiment. Vice versa, the participants who were asked to stand before were now invited to sit down. Once again, first 4 different videos with statements were viewed. Next the height selections were done. Then the small interview took place. After the end of the second phase, the participant was asked to take off the headset and a presence questionnaire and an immersive tendency questionnaire were given to gather related data. Finally, a small interview with one question took place to finish the whole experiment session: *Do you have any comments or thoughts regarding the whole experiment?*

After the post-interview, the participant was thanked and given a movie ticket as the compensation. On average, a session, consisting of the background questionnaire, the tasks, the presence questionnaire, the immersive tendency questionnaire and the small semi-structured interviews took about 60 min. The experiment procedure script can be found in Appendix A.

6.3 Data Collection and Analysis

The author gathered general information from all the participants through the background questionnaire, including age, gender, height and the experiences regarding the HMD and ODV. Five 7-point Likert scale statements, where 1 corresponds to “Totally disagree” and number 7 corresponds to “Totally agree”, two 9-point Likert scale Emofaces, where 1 corresponds to “Totally disagree” and number 9 corresponds to “Totally agree”, were used to evaluate the user experience. Questions regarding the participants’ experience during the use of the software were asked after the sitting and standing conditions respectively, and at the end of the session.

The presence questionnaire and immersive tendency questionnaire were used to collect participants’ subjective feeling. These two questionnaires were originally created by Witmer and

Singer (1998). Some items in the questionnaires were removed because they did not fit the study objectives. Appendix E shows the revised presence questionnaire, appendix F shows the revised immersive tendency questionnaire. At last, the participants' interactions with timestamps were logged in the system. For instance, the activation of the trigger when spotting the icon 'x', changes from one video to another, and video numbers of the selected optimal heights were recorded.

7. Results

The research objectives of the thesis study are to evaluate the effect of the omnidirectional camera height on user experiences, and to evaluate how can sitting and standing viewing position influence the user experience. The result of the user study is demonstrated in this chapter.

7.1 Height Selection

As can be seen in the demographic table in section 6.1, it is easy to notice that the heights of the participants are range from 154cm to 186cm. In order to statistically analyze the influence of camera height and user height relation on user experiences, the participants were categorized into two different groups based on their height: the shorter height group, participants between 150cm and 169cm were divided into this group; and the taller height group, consists of the participants between 170cm to 190cm. As a result, there were 9 participants in the shorter height group and 7 participants in the taller height group. Specific group information can be seen in table 7.1 and table 7.2.

Participant Number	11	14	13	12	16	3	5	2	1
Height	160cm	165cm	163cm	165cm	168cm	162cm	160cm	154cm	156cm

Table 7.1 participants of the shorter height group

Participant Number	15	9	10	4	8	7	6
Height	178cm	175cm	183cm	178cm	182cm	186cm	183cm

Table 7.2 participants of the taller height group

During the height selection stage of the experiments 4 different experienced height types were logged: (1) Most natural one; (2) Most comfortable one, this one was explained as the most enjoyable one while doing the experiments; (3) Highest acceptable one; (4) Lowest acceptable one. The 16 videos used in the height selection part were numbered from 1 to 16, 1 represents the highest height and the 16 represents the lowest height, Table 7.3 shows the correspondence between the height and the video number. Notice that these heights are the height of the camera lens. Each type contains sitting and standing conditions, and for each condition, the data was divided into taller height group (represent with T) and shorter height group (represent with S) respectively. The SPSS software was used to analyze the data, independent samples T-Test was applied to every condition to check if there was any significance difference between the conditions, the significant value for the T-Test was 0.05. One thing needs to be mentioned is that there were 4 missing data due to the unexpected software crash.

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Height(cm)	219	209	199	189	179	169	159	149	139	129	119	109	99	89	79	71.5

Table 7.3 The correspondence between the heights (the height of camera lens) and video numbers

7.1.1 Most Natural Height

For the *Most Natural Height* selection, in the case of viewing while sitting, the average video number of the shorter height group (150cm to 169cm) was 12.13, which corresponds to 108cm, with a standard deviation (SD) of .835 and standard error mean (SEM) of .295. And the average video number of the taller height group (170cm to 190cm) was 10.71, corresponds to 122cm, with SD of .756 and SEM of .286. The average experienced height difference between shorter group and the taller group was around 14cm.

In the case of viewing while standing, the average number for the most natural height of the group S was 8.78, around 141cm, with SD of .667 and SEM of .222. The average number of the group T was 6.86, around 160cm, with SD of .690 and SEM of .261. The average experienced height difference between group S and group T was around 20cm. The overall data group statistics can be seen in Table 7.4 and Figure 7.1 shows a more intuitive bar graph.

Most natural height selection for sitting and standing conditions					
	group	N	Mean	Std. Deviation	Std. Error Mean
sit	S	8	12.13	.835	.295
	T	7	10.71	.756	.286
stand	S	9	8.78	.667	.222
	T	7	6.86	.690	.261

Table 7.4 Most natural height selection group statistics

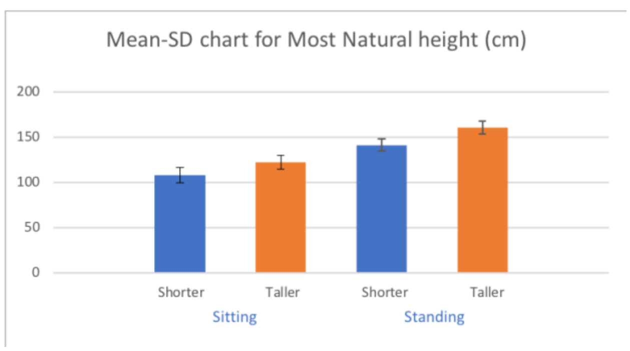


Figure 7.1 Average camera heights with SD for most natural height selection

To verify if there was a significant difference between the data from group S and group T, an independent samples T-test was conducted based on the data and part of the results is shown in Table 7.5. As the results confirmed, sitting condition: $t(15) = 3.411$, $p < 0.05$; standing condition: $t(16) = 5.631$, $p < 0.05$. For both sitting and standing, statistically significant differences did exist between the group S and group T. This means that the camera height of the ODV does influence the participants' feeling. The ODV can transfer an experience similar to the reality to the participants. The result shows that for group S, the average natural camera height for sitting was around 108cm (about 52cm ±10cm lower than the participants' actual height) and around 141cm (about 19cm ± 10cm lower than the participants' actual height) for standing. For group T, the average natural camera height while sitting was about 122cm (around 58cm ± 10cm below their actual height) and about 160cm (around 20cm ± 10cm below their actual height) for standing.

The result shows that for the most natural height selection, the height difference between camera height and participants' actual height is around 20cm: 141cm camera height to 160cm actual height for group S, and 160cm camera height to 180cm actual height for group T. Considering that the camera height is the camera lens height, while the height of the participants is calculated from the top of the head, 20cm difference is normal and acceptable.

Independent Samples T-Test						
	Levene's Test for Equality of Variances			t-test for Equality of Means		
	group	F	Sig.	t	df	Sig. (2 tailed)
sit	Equal Variances assumed	.042	.841	3.411	13	.005
	Equal Variances not assumed			3.435	12.974	.004
stand	Equal Variances assumed	.020	.890	5.631	14	.000
	Equal Variances not assumed			5.605	12.810	.000

Table 7.5 Most natural height selection independent samples test, the comparison is between shorter and taller group under sitting and standing two viewing conditions

7.1.2 Most Comfortable Height

For the *Most Comfortable Height* selection, explained as the most enjoyable height in the experiments, the group statistics is shown in Table 7.6. As the table shows, while sitting, the average comfortable height for group S was around 125cm (Mean = 10.38 and SD = 2.326) and around 148cm (Mean = 8.14, SD = 3.237) for group T. This means that the average height difference was about 23cm.

While standing, the average height for group S was around 158cm (M = 7.11, SD = 2.472) and around 174cm (M = 5.43, SD = 2.225) for group T. The height difference was 16cm. The graphical representation of the relationship of the mean and SD is shown in Figure 7.2.

Most comfortable height selection for sitting and standing conditions					
	group	N	Mean	Std. Deviation	Std. Error Mean
sit	S	8	10.38	2.326	.822
	T	7	8.14	3.237	1.223
stand	S	9	7.11	2.472	.824
	T	7	5.43	2.225	.841

Table 7.6 Most comfortable height selection group statistics

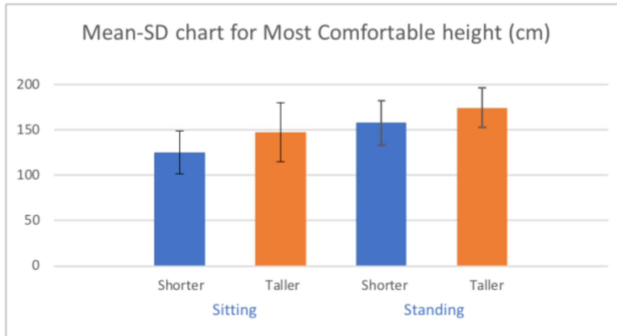


Figure 7.2 Average camera heights with SD for most comfortable height selection

Regarding the most comfortable height selection result, independent T-test was used to verify if there is a statistically significant between group S and group T. The result (in Table 7.7) shows that there was no statistically significant between the data. This interesting result shows that for some of the participants, the most natural height does not exactly mean the most comfortable height. During the experiments, 2 participants mentioned that the most comfortable height should be the one that makes them feel the power of dominance. 1 participant said videos with higher camera height brings her the feeling of sailing, so the higher the height, the more enjoyable she is.

Independent Samples T-Test						
	Levene's Test for Equality of Variances			t-test for Equality of Means		
	group	F	Sig.	t	df	Sig. (2 tailed)
sit	Equal Variances assumed	.097	.760	1.549	13	.145
	Equal Variances not assumed			1.514	10.764	.159
stand	Equal Variances assumed	.008	.932	1.409	14	.181
	Equal Variances not assumed			1.429	13.629	.176

Table 7.7 Most comfortable height selection independent samples test, the comparison is between shorter and taller group under sitting and standing two viewing conditions

7.1.3 Highest Acceptable Height

For the *Highest Acceptable Height* selection, the group statistics can be seen in Table 7.8. In the case of sitting, the average highest acceptable height of the group S, around 141cm (M = 8.75, SD = 2.605) was clearly lower than the group T, around 180cm (M = 4.86, SD = 3.891), the difference of the average height was much higher than the previous two height selection types, about 39cm.

For the standing case, the average highest acceptable height of the group S, approximately 160cm (M = 6.89, SD = 2.315) was lower than the group T, about 189cm (M = 4.00, SD = 2.000) as well, the height difference was around 29cm. Figure 7.3 shows the histogram of the data.

Highest acceptable height selection for sitting and standing conditions					
	group	N	Mean	Std. Deviation	Std. Error Mean
sit	S	8	8.75	2.605	.921
	T	7	4.86	3.891	1.471
stand	S	9	6.89	2.315	.772
	T	7	4.00	2.000	.756

Table 7.8 Highest acceptable height selection group statistics

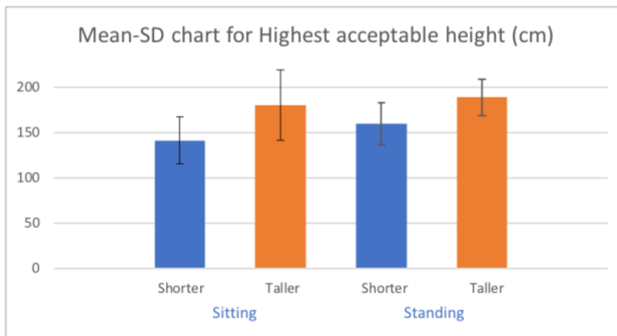


Figure 7.3 Average camera heights with SD for highest acceptable height

Once again, independent samples test was conducted to verify if statistically there is significant difference between group S and T, Table 7.9 shows that the statistically significant differences were detected (sitting: $t(15) = 2.306$, $p < 0.05$; standing: $t(16) = 2.623$, $p < 0.05$). The reason that causes the bigger average video height difference compared to other height selection types could be that there were 3 participants from group T who mentioned that their tolerance for the upper limit of the height is very high. 2 of them said they can have greater feeling of dominance when they feel they are higher.

Independent Samples T-Test						
	Levene's Test for Equality of Variances			t-test for Equality of Means		
	group	F	Sig.	t	df	Sig. (2 tailed)
sit	Equal Variances assumed	3.369	.089	2.306	13	.038
	Equal Variances not assumed			2.243	10.274	.048
stand	Equal Variances assumed	.158	.697	2.623	14	.020
	Equal Variances not assumed			2.674	13.790	.018

Table 7.9 Highest acceptable height selection independent samples test, the comparison is between shorter and taller group under sitting and standing two viewing conditions

7.1.4 Lowest Acceptable Height

For the *Lowest Acceptable Height* selection, the average height difference was rather small. As table 7.10 shows, while sitting, participants from group S's average lowest acceptable height was about 88cm (M = 14.13, SD = .991) and for group T 89cm (M = 14.00, SD = 2.236). There is only about 1cm difference.

While standing, the average lowest height for group S was around 126cm (M = 10.33, SD = 2.598) and around 135cm for group T (M = 9.43, SD = 3.409). The height difference was 9cm. Graphical representation of the means can be seen in Figure 7.4.

Lowest acceptable height selection for sitting and standing conditions					
	group	N	Mean	Std. Deviation	Std. Error Mean
sit	S	8	14.13	.991	.350
	T	7	14.00	2.236	.845
stand	S	9	10.33	2.598	.866
	T	7	9.43	3.409	1.288

Table 7.10 Lowest acceptable height selection group statistics

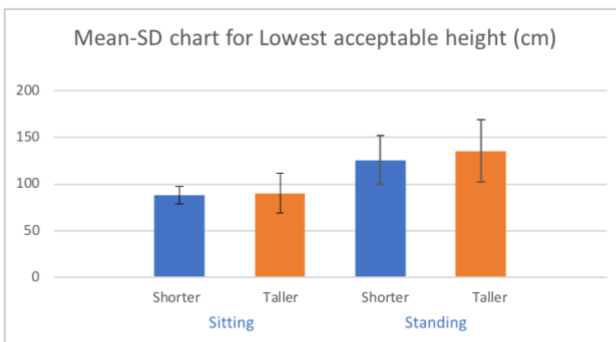


Figure 7.4 Average camera heights with SD for Lowest acceptable height

The statistically significant samples test result (show in Table 7.11) shows that there were no significant differences between group S and T in neither sitting nor standing conditions.

Independent Samples T-Test						
	Levene's Test for Equality of Variances			t-test for Equality of Means		
	group	F	Sig.	t	df	Sig. (2 tailed)
sit	Equal Variances assumed	16.413	.001	.143	13	.888
	Equal Variances not assumed			.137	8.036	.895

stand	Equal Variances assumed	.808	.384	.604	14	.556
	Equal Variances not assumed			.583	10.968	.572

Table 7.11 Lowest acceptable height selection independent samples test, the comparison is between shorter and taller group under sitting and standing two viewing conditions

7.2 User Experience Statements

As mentioned in section 5.1, in total 4 height videos were used for evaluating presence experience with camera heights of 105cm, 132cm, 159cm and 185cm, respectively. During the experiments, videos that correspond to these four heights were labeled from V1 to V4 in the software. The correspondence relationship can be seen in Table 7.12.

This section presents the summary of the user experiences statements responses for different conditions. Similar to the previous section, in this section, data gathered from the participants was analyzed based on two conditions: sitting and standing, and each condition contains two groups: S (participants' height range from 150cm to 169cm) and T (participants' height range from 170cm to 190cm).

Number	V1	V2	V3	V4
Height(cm)	105	132	159	185

Table 7.12 The correspondence between the heights (camera lens height) and video numbers

7.2.1 Sitting Condition: Shorter Group

Figure 7.5 shows the summary of group S's Likert scale statements responses while sitting. As the results show, both V1 and V2 can provide the better feeling of 'being there' compared with other two videos. Compared with V2, participants felt more comfortable with V1, no one felt uncomfortable while experiencing this height environment. Some of the participants mentioned that they felt actually sitting there when the video height felt close to their actual sitting height, thus they tended to feel more 'being there' and comfortable under this circumstance. This was also reflected from the statement '*I felt my experiences in the environment seem consistent with my real-world experience*' and the *Pleasure Emoface*, for which the average responses showed positive results for V1.

It seems that the feeling of how easy it is to look around is not influenced by the height. The participants mentioned this feeling was more related to the technical issues such as the quality of the videos or if they could freely turn around. For the statement '*I felt the objects in the environment appear geometrically correct*', interesting actions were observed during the experiments. Some participants tried to walk around the room to check if there was a distance change and to compare the height with the actors in the videos, showing that they already had a feeling of getting immersed and 'being there'. The average responses of the *Dominance Emoface* indicate that generally the higher camera position can bring more feeling of mastery. Few participants mentioned they can gain a so-called '*god's perspective*' feeling while watching from a higher viewpoint. However, this is not absolute. One participant said that his/her feeling of mastery was lowered because no one was holding an ‘x’, and another participant decreased the score due to the slight acrophobia.

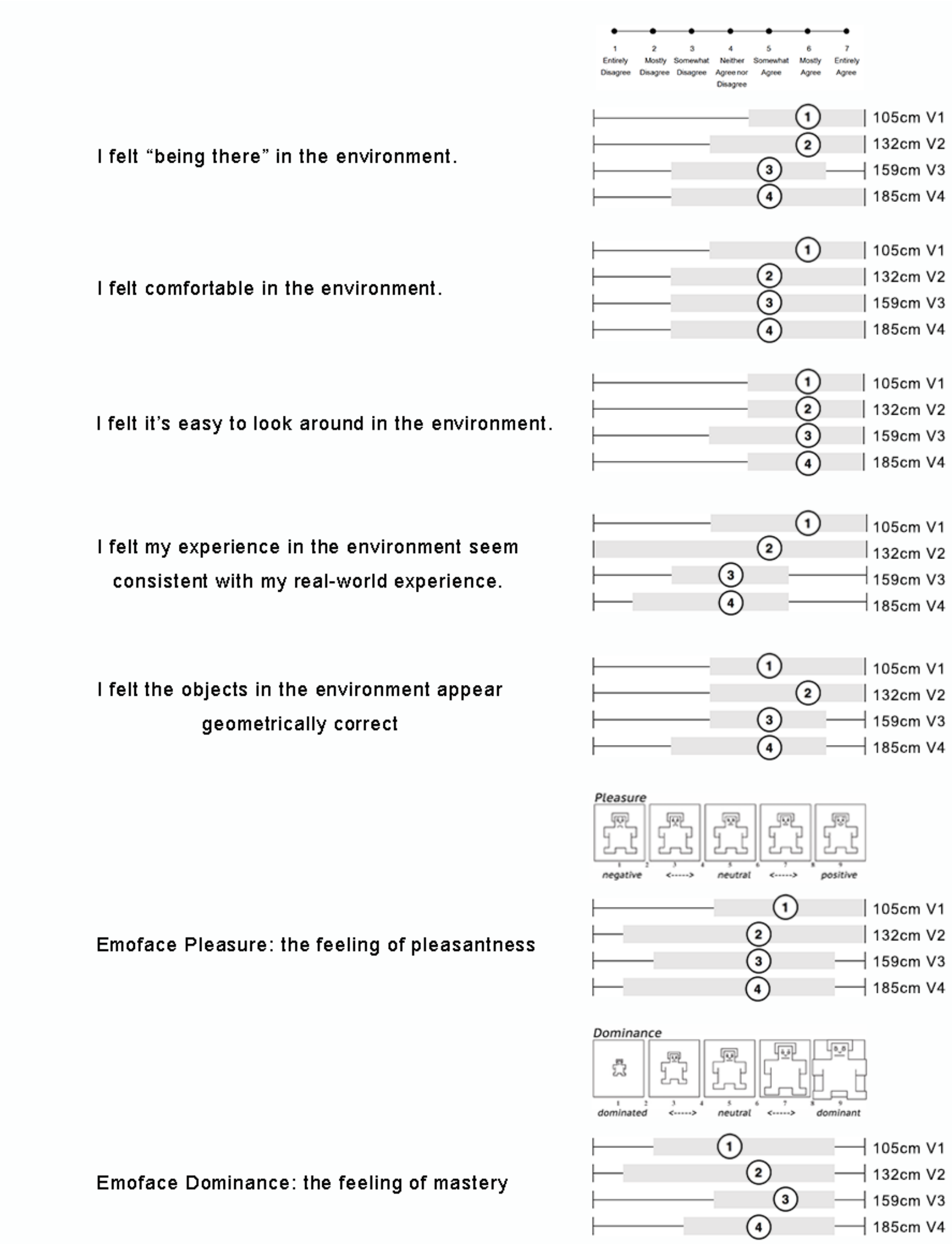


Figure 7.5 User experiences responses summary of group S (150cm-169cm) while sitting. Numbered circles represent the average experience values of the videos with different heights (n = 9): (1) V1, 105cm; (2) V2, 132cm; (3) V3, 159cm; (4) V4, 185cm. Grey boxes represent the participants’ selection range. The whiskers represent the maximum and the minimum responses of the statements.

7.2.2 Sitting Condition: Taller Group

Figure 7.6 shows the summary of group T's Likert scale statements responses while sitting. The results indicate that to some extent all the videos had a positive influence on 'being there' on average, especially V1 and V2, the experience responses from these two videos were all above the middle level (Neither agree nor disagree). When connected with the statement '*I felt comfortable in the environmen*t', V2 showed a relatively high rating: the comfortable experiences only ranged in '*Mostly agree*' and '*Entirely agree*'. V2 had the highest average responses on the statement '*I felt my experiences in the environment seem consistent with my real-world experience*' as well. Overall these results showed that V2 provided the best experience compared to the other videos.

For the statement '*I felt the objects in the environment appear geometrically correct*', the participants seemed to have different judgment references. Some participants used the ceiling as the reference, some used desks or compared themselves with actors in the videos, and some combined the two together. Thus, the result of this statement can be used as a reference, but it contains many personal factors.

The result of the *pleasantness Emoface* needs to be combined with the result of the statement '*I felt my experiences in the environment seem consistent with my real-world experience*'. From the result we can notice that V3 had a wide range response (from lowest to highest, only V2 and V3 contains the highest result) on the consistent statement, which indicates some participants from group T felt the V3 was the most suitable height for this statement. To some extent this influence the result of the *pleasantness Emoface* since those participants gave high score on this one. This problem occurs due to fewer participants, but on the other hand, this also reflects the users' sensitivity to camera heights.

Responses summary result of *Dominance Emoface* showed a relatively stable state. The feeling of mastery gradually increases as the camera height increases.

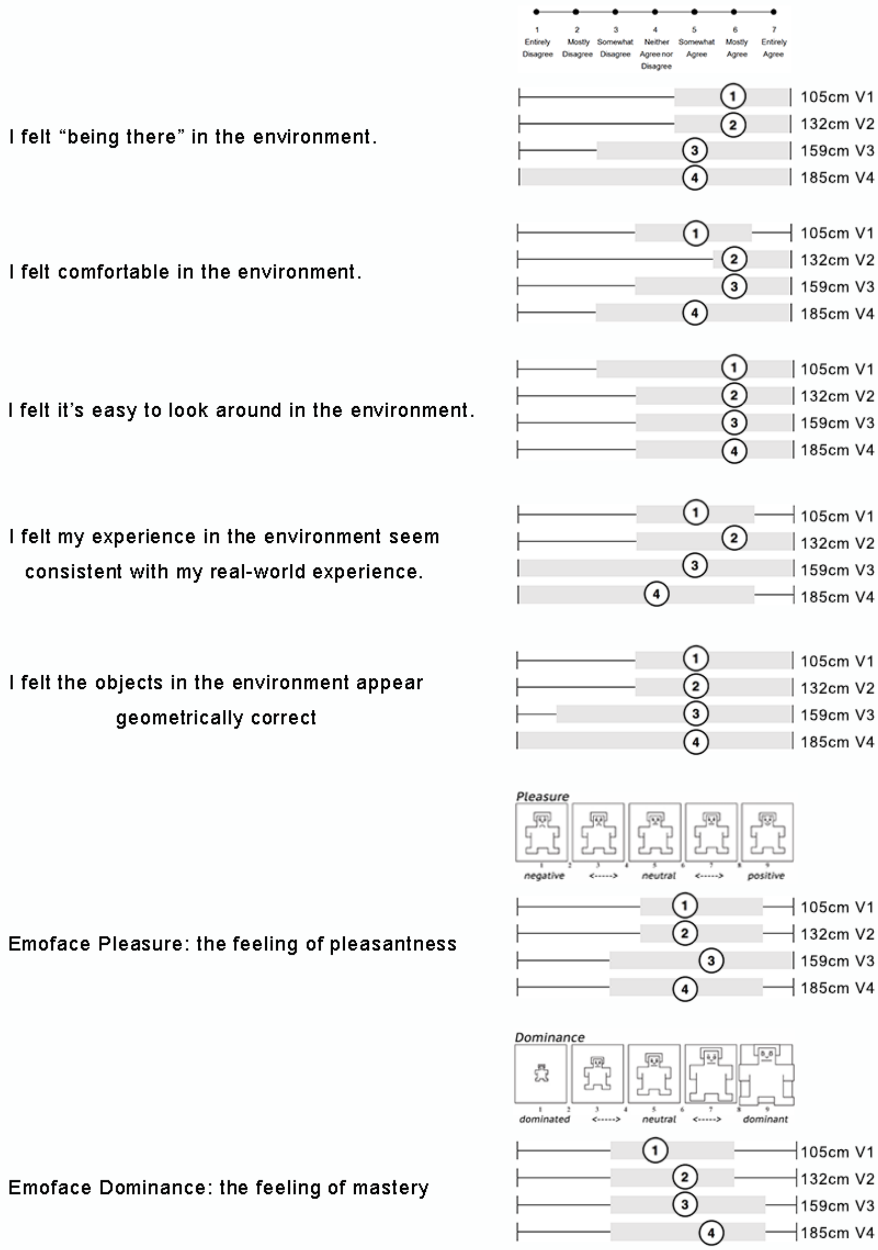


Figure 7.6 User experiences responses summary of group T (170cm-190cm) while sitting.

7.2.3 Standing Condition: Shorter Group

Figure 7.7 shows the summary of group S’s Likert scale statements responses in the case of standing viewing conditions. Results indicated that V2 was the video that engaged participants the most and the one that was closest to participants’ real height experiences. As V2 had the highest average responses on statements ‘*I felt being there in the environment*’ and ‘*I felt my experiences in the environment seem consistent with my real-world experience*’. Both V2 and V3 had a positive influence on the participants’ feeling of comfort.

Participants' average pleasure levels were pretty average, only V1 got a bit lower average response. Still, the average dominance level showed a trend of gradually increasing as the video height increased.

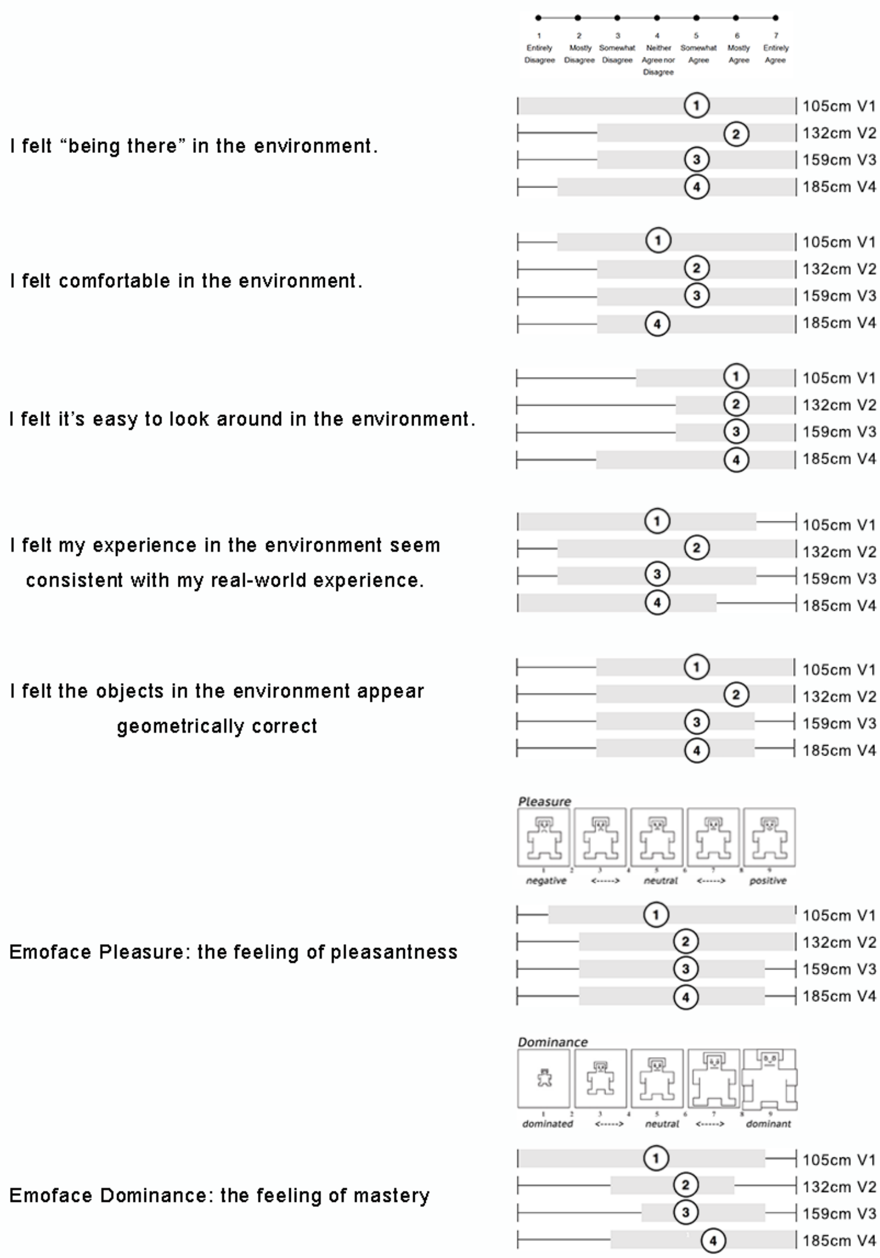


Figure 7.7 User experiences responses summary of group S (150cm-169cm) while standing.

7.2.4 Standing Condition: Taller Group

In Figure 7.8, the summary of group T's Likert scale statements responses in the standing viewing condition can be seen. The result shows a bigger difference among videos compared to the previous conditions. While experiencing the V3, the participants had the best feeling of 'being there'. V1 got negative results both for the statements '*I felt being there in the environment*' and '*I felt my experiences in the environment seem consistent with my real-world experience'*, while V3 got the best average results for both. The average pleasure level was quite positive for V3 as well. Overall as a conclusion, V3 had the most acceptable height for group T while standing.

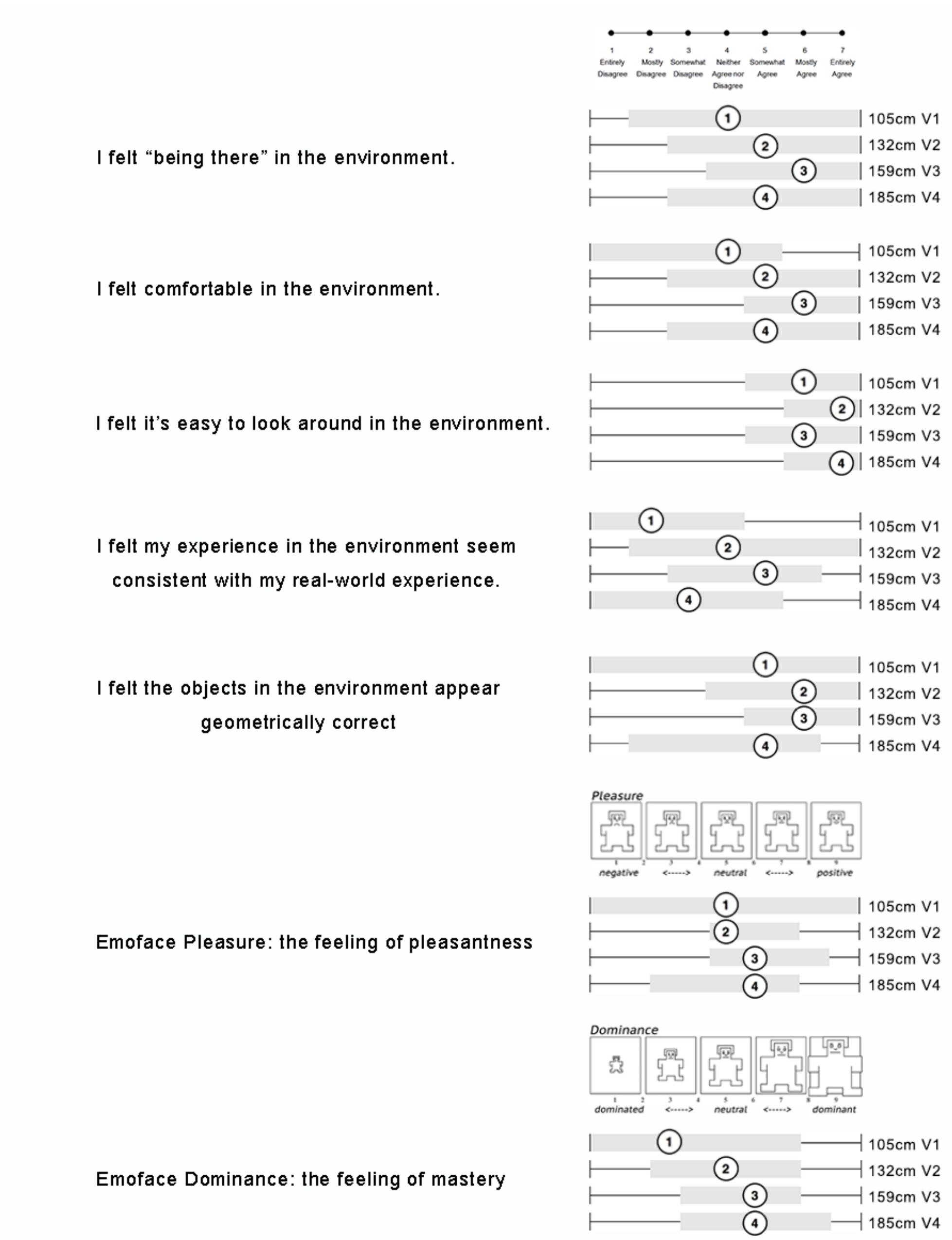


Figure 7.8 User experiences responses summary of group T (170cm-190cm) while standing.

7.3 Interviews

For every experiment, three small semi-structured interviews were conducted. One after sitting condition, one after standing condition and one was at the end of the entire experiment session.

Interviews were recorded as audio recordings and later were transcribed into texts. Comments and suggestions that convey the same feeling or meaning were grouped together and are discussed in this section. Some feedbacks that was generated during the task procedures is included in this section as well.

Following questions were asked in first two interviews:

- *How are you feeling now?*
- *Do you have any feelings regarding the four different height videos?*
- *Compared with sitting/standing, do you have different feelings? (asked in the second interview)*
- *Did you feel dizzy or uncomfortable at any point? If yes, what caused it based on your understanding?*

And one question was asked at the end of the experiment:

- *For the whole experiment, including task procedures, technical parts etc., do you have any comments or suggestions?*

Further questions were asked based on participants' performance or responses during the interviews. The questions were more about the participants' overall feeling and feelings regarding standing and sitting, and to investigate motion sickness in ODV-based virtual environment. To better present the summary, the 16 participants have been labeled from P1 to P16.

7.3.1 Feedback on the Experiment

All the participants thought the experiments gave interesting experiences, P1, P8, P12, P14, and P16 showed curiosity and excitement of using HMD and watching ODVs, P8 said: "*it is like the way to see the future.*". P5 and P14 mentioned the feeling of seeing how the taller people's world looks like was really cool, and it is great to experience their dream height.

As the content of the videos used in the experiments were simple and repetitive, there were seven participants (P3, P5, P9, P10, P11, P12, and P13) who suggested to enrich the video content. Participants P3, P11, and P12 said they were a bit confused at the beginning of the tasks, so they suggested to add some practice videos to help participants to better understand the tasks. Participants P5, P9, P10, and P13 thought the videos could contain more interesting or colorful contents, which can help reduce the visual fatigue and provide better visual experience.

7.3.2 Influence of Camera Height

In total eight participants (P1, P2, P3, P4, P5, P6, P13, and P16) showed the dislike attitude for the highest height video. It seems that when the height was too big, the participants tended to feel it was unreal and scary and there was a slight distortion of the content. The related comments were "*I did not feel like my feet were on the ground*", "*It is a bit scary, I felt like I would fall*", "*I felt I was in the air, this was inconsistent with my actual experience*", "*I felt a bit dizzy*". Among these participants, P1, P3, P4, P6, and P16 said both the highest and the lowest height made them feel uncomfortable and to more easily to feel dizziness.

For the lowest position, P1 said it felt like the actors in the videos may scratch her while walking and P6 said the sense of control was reduced. P8 said lower heights made him feel uncomfortable and he felt a bit dizzy when the height reached the lowest.

There were 7 participants (P2, P3, P4, P5, P12, P13, and P16) who mentioned they tend to feel more comfortable when the experienced height was similar to their actual height.

7.3.3 User Experiences

Some objective factors that influenced the user experiences were mentioned by the participants. Six participants (P3, P4, P10, P12, P13, P16) said the resolution of the videos was a bit low, thus reduced their feeling of getting immersed or engaged. Seven participants (P2, P6, P9, P10, P11, P12, P13) complained about the heavy headset. Participants P1, P4, P6 and P12 suggested adding sounds or music in the videos, as it may help them transfer part of the attention to the audio parts, instead of concentrating on visual parts or the silence environment.

Subjective feedbacks on what affected the user experiences was proposed as well. Participants P8, P10, P12, P13 and P16 mentioned that their feeling of realism or involvement was reduced due to the lack of some objects or user embodiments: chair (while sitting), shadow, feet or even the front hair. Four participants (P3, P4, P7, and P12) tried to walk around during the tasks to check if the distance in the videos were changed. P1 and P2 said that it was a bit disappointing that they cannot interact with the videos. This to some extent lower the feeling of involvement.

7.3.4 Sitting and Standing

Regarding the experiences of sitting and standing, participants P2, P3, P4, P5, P6, P8, P12, and P15 clearly expressed their preference for standing. P2 and P3 said they felt less dizzy while standing. P3 commented that standing felt more 'being there' and realistic. P4, P5 and P8 said that the feeling of sitting was not as real as it could have been in the virtual environment since there was no chair in the videos and that standing felt more natural. P6 and P8 mentioned it took less effort to look around while standing, as there was no need to use force to rotate the chair. On the other hand, standing position did bring an unsafe feeling to some extent. P1, P2, P3, and P10 mentioned there was a feeling of falling while standing, this may cause a little panic.

Participant P1, P7and P14 claimed they prefer sitting position because it felt safer while sitting, as there was no need to worry about the feeling of falling. P13 said since she cannot move while sitting, it felt more realistic while sitting.

Participants P9, P11 and P12 mentioned that compared with standing, the acceptable level of different heights was higher while sitting. The participants were more sensitive to camera heights in the case of standing.

7.3.5 Motion Sickness

In total seven participants (P1, P2, P3, P4, P13, and P15) felt a bit dizzy while sitting and four participants (P1, P4, P7, and P9) felt a bit dizzy while standing. The main causes of dizziness were reported to be the followings: (1) Rotating while looking around or moving around (P1, P2, P3, P9, P13, P14, P15); (2) The quality of the videos, blurred edges (P2, P3, P4, P12, P15); (3) Heavy headset (P1); (4) Camera height was too high (P1, P4); (5) Looking up while camera height was too low (P4, P8).

7.4 Presence Questionnaire and Immersive Tendency Questionnaire

During the experiments, a presence questionnaire (PQ) and an immersive tendency questionnaire (ITQ) were given after both standing and sitting sessions were completed. As mentioned in section 6.3, these two questionnaires were originally developed by Witmer and Singer (1998), and their research showed that the PQ scores were positively correlated with ITQ scores. Which means if the participant had a high score on ITQ, i.e., he/she had a greater possibility to get immersed, then he/she should get a better score on presence questionnaire as well. In order to investigate whether this relationship exists in the ODV-based virtual environment used in the experiments, the results of the two questionnaires are summarized and analyzed in this section.

In the presence questionnaire, there were three questions related to audio factors: question 3, 8 and 9. Since only a short period of sounds existed in the transitional videos, some participants said they did not notice the sounds that much. To make sure the result of these three questions did not interfere with the overall questionnaire results, they were removed from the datasets. For ITQ and PQ, the final scores of each participant (the scores for all the questions from one questionnaire were added together to form the final score) were calculated respectively and formed a one-to-one relationship.

Figure 7.9 shows the ITQ-PQ scatter plot that was created based on the participants ITQ and PQ results. It is easy to see that there is no linear relationship between ITQ and PQ scores. Further Spearman’s correlation coefficient was calculated, and the result can be seen in Table 7.13, the correlation coefficient indicated that there is no significant correlation between ITQ scores and PQ scores. Unfortunately, the conclusion did not match Witmer and Singer’s finding. This may due to the inconsistency and the frequent interruptions while using the software. ODVs with different heights, user experience statements, and transitional videos appeared one after another. The author had to gather feedback after every single ODV, so the participants were always being aware of the existence of the author.

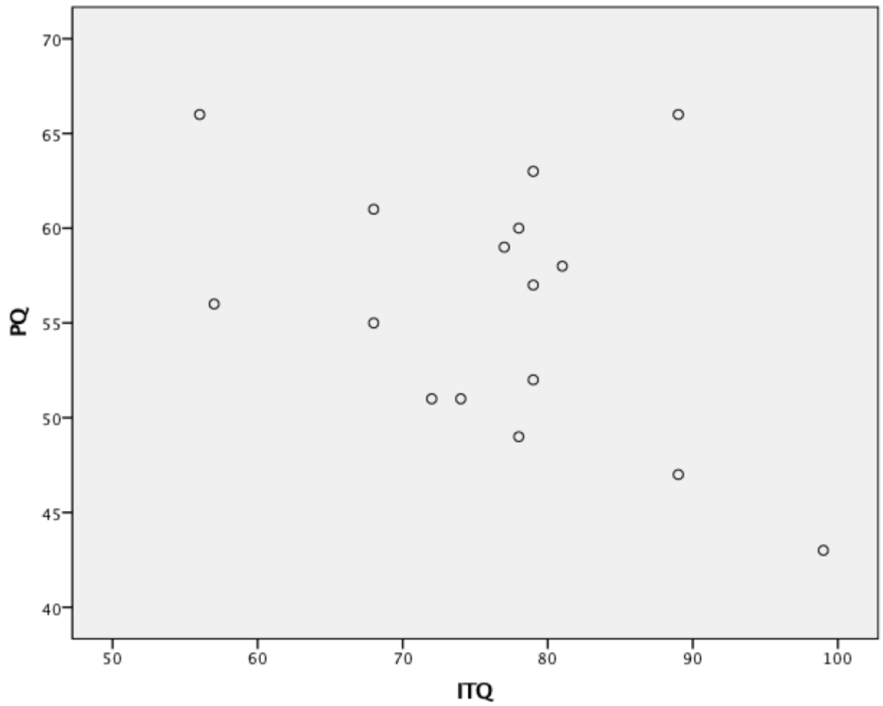


Figure 7.9 ITQ-PQ scatter

Correlations				
			PQ	ITQ
Spearman's rho	PQ	Correlation Coefficient	1.000	- .242
		Sig. (2-tailed)	.	.366
		N	16	16
	ITQ	Correlation Coefficient	- .242	1.000
		Sig. (2-tailed)	.366	.
		N	16	16

Table 7.13 Spearman's correlation for ITQ and PQ questionnaires results

There were many interference factors that can influence the PQ results during the experiments, so the result can only be considered as an attempt but not as the basis of analysis. Other research has shown that ITQ and PQ can have a positive correlation under specific virtual environments (Johns, et al., 2000). Perhaps further experiments can be conducted under specific environments to verify the relationship between the two.

8. Discussion

The findings show that for users, 360-degree videos or virtual reality techniques are novel and worth looking forward to, and it is always an interesting and ‘cool’ experience. However, the experimental procedure and the content of the ODVs were considered a bit simple and not particularly interesting, which to some extent decrease the feeling of being engaged and the visual experience did not reach the participants expectations.

8.1 Camera Heights

The camera height which generates the closest height experience to user’s actual height is considered as the most comfortable and most acceptable camera height, this can be partly explained by user’s prior experiences in real life, the user experiences under this circumstance seems consistent with reality.

In this thesis study, the omnidirectional camera used for ODV recording was Insta360 Pro. For this camera, when we want to record a video that is mainly viewed in a standing position, the suitable height differences between camera height and user’s actual height is considered as 20cm ± 5cm, which can be interpreted as:

User’s actual height – (20cm ± 5cm) = The height of the camera (For videos that will be watched in a standing position)

As I mentioned before in section 7.1.1, considering that the camera height is the position of camera lens, and the user height their total height, the 20cm adjustment naturally corresponds to the distance between user's eye level and total height.

On the other hand, if we want to record a video that is mainly viewed in a sitting position, 55cm ± 5cm was considered as the suitable height differences between camera height and user’s actual height, which means:

User’s actual height – (55cm ± 5cm) = The height of the camera (For videos that will be watched in a sitting position)

Note that this 55cm ± 5cm differences also includes the height from user’s eye to vertex-of-the-head, which is around 20cm. This conclusion means that, Although the height of the chair that was used in the experiments is around 60cm, 35cm ± 5cm differences is considered as the perceptual height differences between sitting and standing positions for users. A possible explanation for this result is that it is very likely that the users could not accurately perceive the height of the chair while watching ODV in a virtual environment. Instead, they use their own ‘preferred sitting heights’ as the reference, and this may depend on personal characteristics. However, since the chair used in the experiments was a normal armchair, we can assume that participants perceived themselves in a most natural sitting position (see Figure 8.1). As demonstrated in Figure 8.1, while under sitting position, the height of the participant should exclude the thigh length, and 35cm ± 5cm is normal for the thigh length, which means the height differences is acceptable.

When applied to actual ODV-based applications, the constants of equations above can be adjusted a bit, as it was considered acceptable if the perceived ODV height was a bit higher or lower, in the range from 10cm to 15cm.

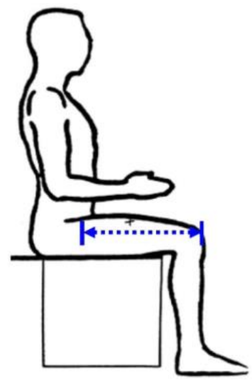


Figure 8.1 Natural sitting position (from NHANES anthropometric manual), blue dot line represents the thigh length. Adapted from Bogin & Varela-Silva (Bogin & Varela-Silva, 2010)

However, too high or too low would result in a bad user experience. The reasons may come from different aspects: an important one is the feeling of inconsistency with actual experience; one participant reported that she has slight acrophobia; several participants mentioned the feeling of falling and dizziness; two participants mentioned the feeling of being so small; when the camera heights were too low, actors in the ODV looked bigger and one participant said this brings the feeling of getting ‘*scratched*’, and the feeling of dominance can be highly decreased.

The subjective feelings regarding sitting and standing depend on many aspects. First, the content of ODV and the embodiments in the virtual environment will greatly influence the user’s judgments. As actors in the ODV are walking around and the chair was not visualized in the ODV, users are very likely to think they ‘should also be standing in the video’, so they may consider standing as a proper position. However, another participant mentioned that since user’s movements were not reflected in the ODV, sitting felt more realistic.

Two participants mentioned that they felt less dizzy while standing. The reason may relate to users’ interaction method: compare to standing, sitting in a chair and rotating seems easier and more convenient, but this also results in a problem: users may spin too much or too fast. Two other participants presented a different point of view: standing takes less effort to look around.

On the other hand, standing also brings the feeling of being in the air, which may cause the feeling of panic, as one participant said she was afraid of moving as it felt like she was going to fall. Sitting in a chair can brings the feeling of ‘having a base’ and reduce the feeling of falling, which is safer. This sense of safety can be the reason for higher acceptable level of different heights while sitting.

8.2 Immersion and Presence

The experiment finding suggests that ODV-based software with HMD do bring a high-level presence and immersive experiences. The different heights may influence these experiences to some extent, but overall the presence experience, or the feeling of 'being there' is positive. For the immersive experience, some of the participants' subconscious behaviors indicated that they were getting immersed to some extent. For instance, several participants tried to walk around and interact with the ODV content and some participants were afraid of falling. This may have many explanations: first, the HMD effectively blocks other visual disturbances; besides, the HMD well-matched the features of ODV: it provides a large field-of-view and a three-dimensional viewing experience; finally, the built-in head tracking of HMD ensures a consistent and fluent viewing experience.

Several aspects that break the immersion and presence were revealed in findings. Six participants said the video quality reduced the feeling of being engaged, as they can immediately realize the difference between the ODV and the real world after taking off the HMD. Seven participants mentioned the heavy headset broke the feeling of immersion, as they can constantly realize that they are wearing the VR headset. Four participants reported the lack of sound modality reduced the immersive experience. Five participants informed the missing of embodiments in the virtual environments, which include chair, feet or shadow and two participants said that the fact that they could not interact with the ODV was a bit disappointing. These can be addressed by combining OVD with other VR techniques. Four participants tried to move around but failed to see the corresponding changes in the ODV which decrease the feeling of involvement. There were at least three participants who mentioned that the inconsistent content in the video made them confused and broke the feeling of immersion to some extent. This was due to the looping videos, when a video ended, the content of the video will reset to the beginning which caused sudden changes.

In total eight participants reported dizziness during the experiments. Most of them felt dizzy while rotating the chair and saw the blurred edges of the ODV, this is mostly due to the technical problems and can be addressed with well-developed techniques. Two participants felt a bit dizzy when they were experiencing the highest camera height video. This result may due to the acrophobia as one of the participants mentioned she has a slight acrophobia. Two other participants felt dizzy when the video height reached the lowest level. This may be caused by the resistance to the current environment: the participants felt they were 'minified' and lost the power of control.

8.3 Future Work

Despite the potential, ODV-based applications still need to be properly designed. In order to improve the feeling of immersion, the ODV content should be carefully designed. If the videos need looping, the end of the video should be managed to link up with the beginning of the video. In the future, the sense of embodiment in virtual reality can be combined with the ODV, the embodiments can greatly increase the sense of presence and help users know what they are doing, and how they should act. Further studies can be done in this field. Another hot topic, interactive ODV application seems to have the potential to provide better immersive experiences as well. Although interaction is not within the scope of this thesis, many related studies already exists. Follow-up studies related to immersion and presence evaluation can be done in this field.

For future ODV-related experiments which include different camera heights, well-designed practice videos should be considered to apply. Including different kinds of height information in the practice videos may help the participants understand what is going to happen and prepare themselves beforehand. Otherwise, users may get confused and give uncertain or biased answers in the initial stage of the experiment. And also, if the experiment procedure contains repetitive contents, it is important to include interesting elements and small breaks to reduce the feeling of getting tired and visual fatigue. A good visual experience can help increase the feeling of immersion.

While the finding of the thesis shows that the most comfortable video height is the height that is closest to participant’s actual height. This raises other research questions, for instance, what kind of ODV contents should the lower camera height or higher camera height be applied to? What if the intention is to bring the feeling of horror, such as a horror house? Maybe lower the height would be a good choice. Or how about a fighting competition similar to Street Fighter? A higher height may bring a better view and increase the feeling of excitement.

Moreover, it is worth to discover more detailed things about how sitting position influences the user experience, for instance, will the type of the chair or the height of the chair influence the user experience while watching ODV. How acrophobia influences user experiences in the virtual environment can be another interesting topic to study in the future. For the relationship of ITQ and PQ, further research can be done in a more well-designed and variable-controlled virtual environment.

9. Conclusion

The 360-degree video, or omnidirectional video (ODV), has been widely used due to its unique advantages: ODV can provide a strong presence and immersive experience, has superior interactive performance, is easy to generate with existing omnidirectional cameras at rather a low cost.

In this thesis, the author investigated the user experience of an ODV-based software. User experiments including tasks, questionnaires and interviews was conducted. To gain insight on user experiences under different omnidirectional camera heights, on experiences regarding sitting and standing, on the presence and immersive experiences and on symptoms occurred in ODV-based applications, the author collected and analyzed system logs which include height selection data, interview records, and questionnaire data.

The thesis findings indicate that, for ODV-based applications, while a user is watching ODV under a standing position, if the perceived ODV height closest to the user's actual standing height is used, the user tends to get the most comfortable and immersive experience.

However, on the other hand, while a user is watching ODV under a sitting position in a virtual environment, the finding of this study shows that 35cm ± 5cm could be considered as widely accepted height differences compared with standing position, while the actual armchair used in the experiments was around 60cm high. The reason for this may be that it is likely that the user is hard to notice the actual chair height that he/she sits while watching the ODV. Under this circumstance, as shown in Figure 8.1, the user may get the most comfortable and immersive experience under his/her most natural or preferred sitting position, as 35cm ± 5cm could be considered as the normal range for the thigh length. This phenomenon can be studied as an interesting topic in future studies. Note that while setting an omnidirectional camera for video recording, the height of the camera should always subtract the height of the eye to the top-of-the-head, which is around 20cm according to the finding of this study.

The choice between sitting and standing viewing position needs to be considered from different aspects, including the characteristics of the target user group, video contents, the time period of using the application etc. While making ODV-based applications, the ODV content needs to be carefully-designed. With the advantage of providing presence and immersive experience, ODV and HMD can easily cause symptoms such as dizziness as well. Symptoms caused by hardware or technical issues will be solved with the development of the technology. More effort can be put on improving video content, for instance, choosing a suitable camera height, using techniques to stabilize the shaky video, increasing the consistency of the video content etc. The purpose is to provide better user experiences in the future.

As ODV-based applications are becoming a popular and grounded technology, it is necessary to produce further research in this area. For the future work, combining the embodiments in virtual reality with ODV, and the influence of embodiments on presence and immersive experience seem to be interesting topics. Interactive ODV applications are the hot trend in ODV-based applications, many studies related to this topic already exists, but how these applications affect user experiences is still worth working on. And also, besides the natural sitting and standing camera heights, lower and higher camera heights can be applied in different contexts, for example, to understand how acrophobia influence the user experience in virtual environments. These topics can help gain a deeper understanding of user experience (including immersion and presence) in ODV-based applications.

References

Benko, H., & Wilson, A. D. (2010, November 7). Multi-point interactions with immersive omnidirectional visualizations in a dome. *ACM International Conference on Interactive Tabletops and Surfaces*, pp. 19-28.

Benko, H., Wilson, A. D., & Balakrishnan, R. (2008). Sphere: multi-touch interactions on a spherical display. *Proceedings of the 21st annual ACM symposium on User interface software and technology* (pp. 77-86). ACM.

Bevan, N. (2009). Extending quality in use to provide a framework for usability measurement. *International Conference on Human Centered Design* (pp. 13-22). Berlin, Heidelberg: Springer.

Bevan, N. (2009). What is the difference between the purpose of usability and user experience evaluation methods. *Proceedings of the Workshop UXEM*, 1-4.

Bleumers, L., Van den Broeck, W., Lievens, B., & Pierson, J. (2012). Seeing the bigger picture: a user perspective on 360 TV. *Proceedings of the 10th European conference on Interactive tv and video* (pp. 115-124). ACM.

Bogin, B., & Varela-Silva, M. I. (2010). Leg length, body proportion, and health: a review with a note on beauty. *International journal of environmental research and public health 7,* 1047-1075.

De la Torre, F., Vallespi, C., Rybski, P. E., Veloso, M., & Kanade, T. (2005). Omnidirectional video capturing, multiple people tracking and identification for meeting monitoring.

Decock, J., Van Looy, J., Bleumers, L., & Bekaert, P. (2014). The pleasure of being (there?): an explorative study into the effects of presence and identification on the enjoyment of an interactive theatrical performance using omnidirectional video. *AI & society*, 449-459.

Delahoche, L., Pégard, C., Marhic, B., & Vasseur, P. (1997). A navigation system based on an ominidirectional vision sensor. *Intelligent Robots and Systems, 1997. IROS'97., Proceedings of the 1997 IEEE/RSJ International Conference* (pp. 718-724). IEEE.

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. *Information and Software Technology*, 789-817.

Foote, J., & Kimber, D. (2000). Flycam: Practical panoramic video and automatic camera control. *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference* (pp. 1419-1422). IEEE.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 345-352). ACM.

Gaspar, J., Winters, N., & Santos-Victor, J. (2000). Vision-based navigation and environmental representations with an omnidirectional camera. *IEEE Transactions on robotics and automation 16*, 890-898.

Hakulinen, J., Keskinen, T., Mäkelä, V., Saarinen, S., & Turunen, M. (2017). Omnidirectional Video in Museums–Authentic, Immersive and Entertaining. *International Conference on Advances in Computer Entertainment* (pp. 567-587). Cham: Springer.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 641-661.

Johns, C., Nunez, D., Daya, M., Sellars, D., Casanueva, J., & Blake, E. (2000). The interaction between individuals' immersive tendencies and the sensation of presence in a virtual environment. *Virtual Environments* (pp. 65-74). Vienna: Springer.

Kallioniemi, P., Mäkelä, V., Saarinen, S., Turunen, M., Winter, Y., & Istudor, A. (2017). User Experience and Immersion of Interactive Omnidirectional Videos in CAVE Systems and Head-Mounted Displays. *IFIP Conference on Human-Computer Interaction* (pp. 299-318). Springer.

Kasahara, S., Nagai, S., & Rekimoto, J. (2015, June 3). First person omnidirectional video: System design and implications for immersive experience. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 33-42.

Kosinski, D. M., & Kosinski, D. M. (1999). *The artist and the camera: Degas to Picasso.* TX: Dallas Museum of Art. Dallas: Dallas Museum of Art.

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 719-728). ACM.

Lewis, J. R. (2014). Usability: lessons learned… and yet to be learned. *International Journal of Human-Computer Interaction*, 663-684.

Mütterlein, J., & Hess, T. (2017). Immersion, Presence, Interactivity: Towards a Joint Understanding of Factors Influencing Virtual Reality Acceptance and Use.

McMahan, A. (2003). Immersion, engagement and presence. *The video game theory reader*, 86.

McNamara, N., & Kirakowski, J. (2006). Functionality, usability, and user experience: three areas of concern. *interactions*, 26-28.

Moczarny, I. M., De Villiers, M. R., & Van Biljon, J. A. (2012). How can usability contribute to user experience?: a study in the domain of e-commerce. *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference* (pp. 216-225). ACM.

Nayar, S. K. (1997). Catadioptric omnidirectional camera. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference* (pp. 482-488). IEEE.

Nielsen, J. (1995). 10 usability heuristics for user interface design. *Nielsen Norman Group*.

Nilsson, N. C., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human Technology*.

Norman, D. A. (1998). The invisible computer. 1998. *Massachusetts: Cambridge*.

Obrist, M., Roto, V., & Väänänen-Vainio-Mattila, K. (2009). User experience evaluation: do you know which method to use? *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 2763-2766). ACM.

Ochi, D., Kunita, Y., Fujii, K., Kojima, A., Iwaki, S., & Hirose, J. (2014). HMD viewing spherical video streaming system. *Proceedings of the 22nd ACM international conference on Multimedia* (pp. 763-764). ACM.

Ochi, D., Kunita, Y., Kameda, A., Kojima, A., & Iwaki, S. (2015). Live streaming system for omnidirectional video. *Virtual Reality (VR), 2015 IEEE* (pp. 349-350). IEEE.

Oh, S. J., & Hall, E. L. (1987). Guidance of a mobile robot using an omnidirectional vision navigation system. *Mobile Robots II*, 288-301.

Onoe, Y., Yamazawa, K., Takemura, H., & Yokoya, N. (1998). Telepresence by real-time view-dependent image generation from omnidirectional video streams. *Computer Vision and Image Understanding*, 154-165.

Onoe, Y., Yokoya, N., Yamazawa, K., & Takemura, H. (1998). Visual surveillance and monitoring system using an omnidirectional video camera. *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference* (pp. 588-592). IEEE.

Pakkanen, T., Hakulinen, J., Jokela, T., Rakkolainen, I., Kangas, J., Piippo, P., . . . Salmimaa, M. (2017). Interaction with WebVR 360° video player: Comparing three interaction paradigms. *Virtual Reality (VR), 2017 IEEE* (pp. 279-280). IEEE.

Palmisano, S., Mursic, R., & Kim, J. (2017). Vection and cybersickness generated by head-and-display motion in the Oculus Rift. *Displays* , 1-8.

Peri, V. N., & Nayar, S. K. (1997). Generation of perspective and panoramic video from omnidirectional video. *Proc. DARPA Image Understanding Workshop*, 243-245.

Petrie, H., & Bevan, N. (2009). The Evaluation of Accessibility, Usability, and User Experience. 1-16.

Petry, B., & Huber, J. (2015). Towards effective interaction with omnidirectional videos using immersive virtual reality headsets. *Proceedings of the 6th Augmented Human International Conference* (pp. 217-218). ACM.

Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using virtual reality to assess user experience. *Human Factors*, 964-982.

Rusu, C., Rusu, V., Roncagliolo, S., Apablaza, J., & Rusu, V. Z. (2015). User experience evaluations: challenges for newcomers. *International Conference of Design, User Experience, and Usability* (pp. 237-246). Springer.

Saarinen, S., Mäkelä, V., Kallioniemi, P., Hakulinen, J., & Turunen, M. (2017). Guidelines for Designing Interactive Omnidirectional Video Applications. *IFIP Conference on Human-Computer Interaction* (pp. 263-272). Cham: Springer.

Scaramuzza, D. (2012). Omnidirectional camera. *Encyclopedia of Computer Vision, K. Ikeuchi, Ed. Berlin: Springer-Verlag*.

Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, 183-201.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence* , 560-565.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 603-616.

Sun, X., Foote, J., Kimber, D., & Manjunath, B. S. (2005). Region of interest extraction and virtual camera control based on panoramic video capturing. *IEEE Transactions on Multimedia 7*, 981-990.

Turunen, M., Hakulinen, J., Melto, A., Heimonen, T., Laivo, T., & Hella, J. (2009). SUXES-user experience evaluation method for spoken and multimodal interaction. *Tenth Annual Conference of the International Speech Communication Association*.

Väätäjä, H., Koponen, T., & Roto, V. (2009). Developing practical tools for user experience evaluation: a case from mobile news journalism. *European Conference on Cognitive Ergonomics: Designing beyond the Product---Understanding Activity and User Experience in Ubiquitous Environments* (p. 23). VTT Technical Research Centre of Finland.

Vermeeren, A. P., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 521-530). ACM.

Wikipedia. (2017, December). *Louis Daguerre*. Retrieved from Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/Louis_Daguerre

Wikipedia. (2018, April). *Kodak DCS*. Retrieved from Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/Kodak_DCS

Wikipedia. (2018, April). *List of average human height worldwide*. Retrieved from Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/List_of_average_human_height_worldwide

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 225-240.

Yagi, Y., & Kawato, S. (1990). Panorama scene analysis with conic projection. *Intelligent Robots and Systems' 90.'Towards a New Frontier of Applications', Proceedings. IROS'90. IEEE International Workshop* (pp. 181-187). IEEE.

Zoric, G., Barkhuus, L., Engström, A., & Önnevall, E. (2013). Panoramic video: design challenges and implications for content interaction. *Proceedings of the 11th european conference on Interactive TV and video* (pp. 153-162). ACM.

Appendix

Appendix A - User experience for ODV application experiment script

- *Prepare work [30 minutes before test]*
    - *Turn on the light*
    - *Press the 'VR/PC' button to turn on the computer*
    - *Turn on the projectors [press 'on']*
    - *Take out the VR headset*
    - *Open steam*
    - *Enter participant number*
    - *Open test software*
    - *Move VR headset [see if it's detected]*
    - *Turn on VR headset handle [see if it's detected]*
    - *If everything is setup, prepare the paper questionnaires for the participant and wait until the test begin.*
- *Before the usage*
    - Introducing the system/evaluation.
        - *Welcome to this test session! I have invited you here to interact with our 360-degree video application and to give me feedback. The idea is not to test you in any way, I just want feedback about your experience regarding the ODV clips so that the result can benefit future developments.*

        - *The information gathered from you will not be shared or published outside from my research. And please note that your participation is totally voluntary, and you are allowed to stop the test whenever you want without explanation.*

        - *In the test, you will be wearing virtual reality glasses and view some video clips. I will ask for your feedback between the video clips, and at the end of the test I'll ask you to fill in several questionnaires and ask you few questions regarding your overall experience.*

        - *Are you ready to start?*
    - Handing out the informed consent form.
        - *Now I'd like you to read this consent form and sign it. If you have any questions, just ask me.*
        - *Here is your copy of the consent form*
    - Handing out the demographic questionnaire
        - Okay. Now *I would ask you to fill in this demographic questionnaire* [Provide the paper questionnaire]

        - *Your participant number is **__XX__**, please enter it there to the first box. If you have any questions, just ask me.*
- *Usage, i.e., a sequence of multiple video clips with varying parameters*
    - Giving instructions, limitations and an overall "main task" for the user.
        - *Okay. The test consists of two sessions, in the first one you'll be asked to stand/sit and in the second one you'll be asked to sit/stand. In each session, there are totally 4 videos, in the video you will see several people bring a paper with an icon on it. Your main task is to explore these 360-degree videos through VR glasses and find the correct icon. The correct icon is always a cross "x", which means you need to find the paper that with an 'x' on it, and the incorrect ones are circle marks. When you have found the correct icon, just tell me you find it and press the button behind the controller. Then we will continue to the next phase.*
    - Explain the middle tasks, 5 statements and 2 Emofaces after each single video.
        - *Okay. After each single video is being played, you will be asked 7 questions, you can see them here to have a basic understanding. First 5 questions are based on a 7 Likert scale and last 2 are based on a 9 Likert scale.*
        - *Handing out the Likert scale question script.*
    - Fixing the focus, calibration and practice round.
        - *Now you can put the virtual glasses on, and we can do some adjustments so that it feels comfortable and the picture is clear. You can adjust the headset from here [show] and adjust the distance between the lenses here [show].*
        - *[If participant wear glasses] You don't need to take off the glasses.*
        - *Do you see the text clearly? [If not, guide focusing the glasses]*
        - *Okay, now we can continue. I will first put on a video which you can explore freely to test looking around in the virtual environment. Remember that you can also rotate your chair (or turn around). [Put practice round video on]*
        - *When you are ready, we can move to the actual tasks.*
    - The actual usage of the application, i.e., a sequence of trials with different conditions.
        - *Okay, let's start. And as a reminder, your task is to find the correct icon shown as cross "x" from the videos, then tell me and press the button behind the controller. Here we go! [Start the test]*

- For each set of 4 video clips:
  - Wait until the participant reads out loud the correct text, then say "*good*" or something, and press space to move to the next scene. (If wrong icon was activated, say "*I think that was not the correct icon, try finding it somewhere else*" or similar.)
  - After the forth trial, press space again to move to the feedback scene.
  - Only the first time, or when necessary:
    Now I will read to you some statements one by one, and I ask you to give me a number from 1–7 based on how much you agree with the statement based on your last video. You can see the scale on the display, so you don't have to memorize it.
    Number 1 corresponds to "Totally disagree" and number 7 corresponds to "Totally agree". [7 Likert Scale]
    Now here are two types of Emofaces, one represents the feeling of enjoyment and pleasantness, another represents the feeling of mastery, please choose the one that best represents your current feelings.

    If you feel uncomfortable, you can remove the headset for a while.
  - *Now let's move on to the statements.* [Read the statements one by one and mark the answers to the Excel sheet on the row with the correct participant number!]
    - (1.) *I felt "being there" in the environment.*
    - (2.) *I felt comfortable in the environment.*
    - (3.) *I felt it's easy to look around in the environment*
    - (4.) *I felt my experience in the environment seem consistent with my real-world experience.*
    - (5.) *I felt the objects in the environment appear geometrically correct (seem to have the same size and distance in relation to myself and other objects).*
    - (6.) Emofaces [choose one for the current video]
  - After gathering the answers, press space to move to the next condition, i.e., next set of 4 trials in standing condition (or end the test).
- After each set of 4 video clips:
  - Select the optimal video for these four conditions:
    - (1.) Most natural one [press 'N']

            - (2.) Most enjoyable (comfortable) one [press ‘C’]
            - (3.) Lowest acceptable one [press ‘L’]
            - (4.) Highest acceptable one [press ‘H’]
        - *Okay*, now you can remove your headset and take a break before we move to second session*, I will ask you few questions, the conversation will be recorded*:
            - *How is your feeling now?*
            - *Does the environment makes you feel cybersickness (or nausea, dizzy) at some point?* If yes, please specify.

- *After the usage, i.e., after all trials are finished*
    - Gathering overall user experiences with a questionnaire and a simple interview.
        - *Great, you now completed all tasks! You can remove the headset now.*
        - *Now please fill in these two questionnaires.*
            - *Handing out the presence questionnaire*
            - *Handing out the immersive tendency questionnaire.*
        - *After the participant is ready with the questionnaire:*

          Open question in the end:
            - *Do you have any comments or thoughts regarding the whole experiment?*

          Thank you very much for your feedback and participation! From my side, you are now free to leave, unless you have any questions or comments.

Appendix B - Script: Statements that will be asked after each single video

(1) *I felt "being there" in the environment.*

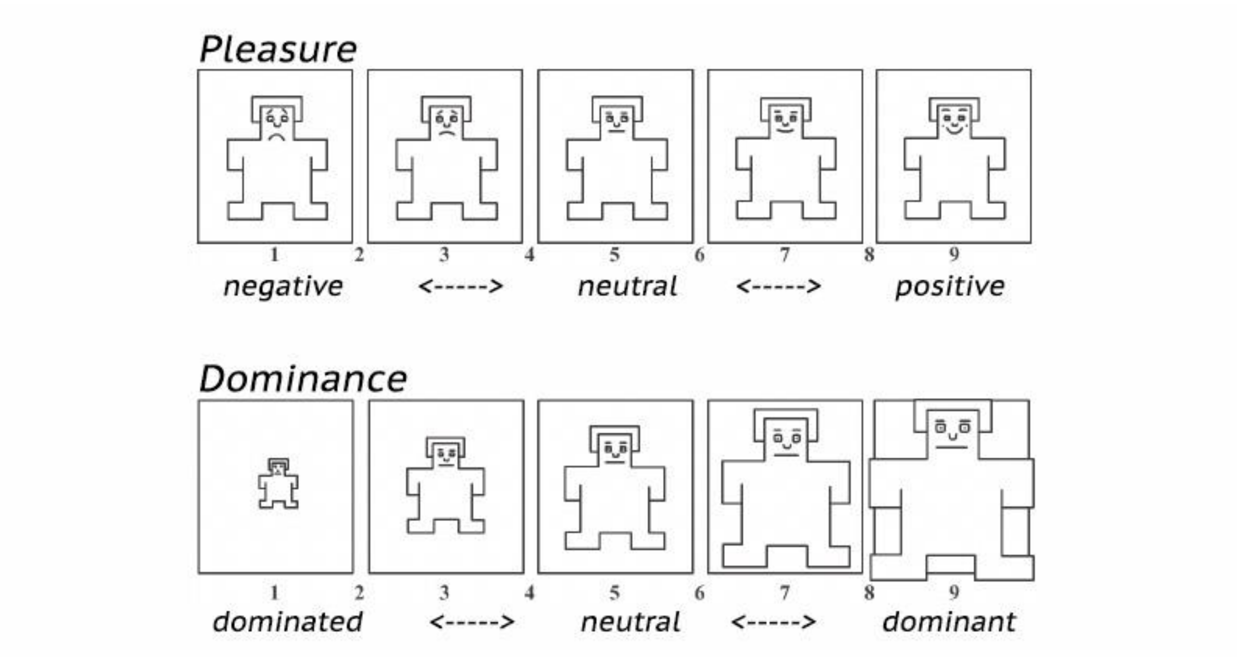
(2) *I felt comfortable in the environment.*

(3) *I felt it's easy to look around in the environment*

(4) *I felt my experience in the environment seem consistent with my real-world experience.*

(5) *I felt the objects in the environment appear geometrically correct (seem to have the same size and distance in relation to myself and other objects).*

(6) *Emofaces:*
***Pleasure** represents the feeling of enjoyment and pleasantness.*
***Dominance** represents the feeling of mastery.*



Appendix C - Consent form

User experience test consent form

Please read and sign this form.

You have been invited to participate in a user experience test which is part of my master's thesis work at the University of Tampere. By participating in the test, you will help to evaluate the user experience of 360-degree video application.

In this user experience test:

- You will be asked to use a head-mounted display to watch several 360-degree video clips.
- You will be asked to watch some of the 360-degree video clips on sitting while others on standing.
- You will be asked to fill in 3 questionnaires.
- You will be asked to answer few questions regarding the video clips you viewed.
- The interview part of the experiment will be recorded as an audio recording.

Participation in this usability study is voluntary. All information will remain strictly confidential. The results and findings may be used to help improve the 360-degree video application. By participating the experiment, you can get a Finnkino movie ticket as a compensation.

You can withdraw your consent to the experiment and stop participation at any time. Feel free to ask any questions you may have about your participation.

If you have any questions after the experiment, please contact Xinru Hu at hu.xinru.x@student.uta.fi

I have read and understood the information on this form and had all of my questions answered

Date and Place: ______________________________

Signature: ______________________________

Name Clarification: ______________________________

Email Address: ______________________________
(The movie ticket will be sent through email)

Thank you!

Appendix D - Background questionnaire

Please take a few minutes to answer the following questions to help me better understand your background. I will use this information only to provide background and usage context in which to interpret the feedback you'll give me in the user study. I will keep your information confidential.

1. Participant number

   ____

2. What is your age?

   ____

3. What is your gender?

   ❑ Male ❑ Female ❑ Other

4. Are you wearing eye glasses at the moment?

   ❑ Yes ❑ No, but I'm wearing contact lenses at the moment ❑ No

5. What is your height?

   ____

6. What is your occupation?

   ____

7. What is your field of study?

   ____

8. Have you ever used VR headset before?

   ❑ Yes ❑ No

   (If 'yes', please answer question 9, if 'no', please go to question 10.)

9. how many times? (e.g. 1-2 times a year, 3-5 times a year, over 5 times a year)

____

Mainly for what purposes?

____

10. Have you ever watched 360-degree videos before?

❑ Yes ❑ No

(If 'yes', please answer question 11, if 'no', please go to question 12.)

11. how many times? (e.g. 1-2 times a year, 3-5 times a year, over 5 times a year)

____

On what platform? (Mobile phone, Personal computer, VR headset or …)

____

Mainly for what purposes?

____

12. Do you had experience of feeling motion sickness, for instance, dizzy or nausea while using some kind of technology before?

❑ No

❑ Yes, for example:

____

____

Appendix E - Presence questionnaire[1]

Indicate your preferred answer by placing an ''X'' in the appropriate box of the scale in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply.

1. How completely were *all* of your senses engaged?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL	MODERATELY COMPELLING	VERY COMPELLING

2. How much did the visual aspects of the environment involve you?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

3. How much did the auditory aspects of the environment involve you?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

4. How aware were you of events occurring in the real world around you?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

5. How aware were you of your display and control devices?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

6. How consistent or connected was the information coming from your various senses?

|_________|_________|_________|_________|_________|_________|_________|

NOT CONSISTENT	MODERATELY CONSISTENT	VERY CONSISTENT

[1] *Original version: Witmer, B.G. & Singer. M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225-240.

7. How much did your experiences in the virtual environment seem consistent with your real-world experiences?

|________|________|________|________|________|________|________|
NOT CONSISTENT — MODERATELY CONSISTENT — VERY CONSISTENT

8. How well could you identify sounds?

|________|________|________|________|________|________|________|
NOT AT ALL — SOMEWHAT — COMPLETELY

9. How well could you localize sounds?

|________|________|________|________|________|________|________|
NOT AT ALL — SOMEWHAT — COMPLETELY

10. To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session?

|________|________|________|________|________|________|________|
NOT AT ALL — SOMEWHAT — COMPLETELY

11. How involved were you in the virtual environment experience?

|________|________|________|________|________|________|________|
NOT INVOLVED — MILDLY INVOLVED — COMPLETELY ENGROSSED

12. How quickly did you adjust to the virtual environment experience?

|________|________|________|________|________|________|________|
NOT AT ALL — SLOWLY — LESS THAN ONE MINUTE

13. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

|________|________|________|________|________|________|________|
NOT AT ALL — INTERFERED SOMEWHAT — PREVENTED TASK PERFORMANCE

14. How much did the control devices interfere with the performance of assigned tasks or with other activities?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL INTERFERED SOMEWHAT INTERFERED GREATLY

15. Were you involved in the experimental task to the extent that you lost track of time?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

Appendix F - Immersive tendency questionnaire[2]

Indicate your preferred answer by placing an ''X'' in the appropriate box of the scale in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. For example, if your response is once or twice, the second box from the left should be marked. If your response is many times but not extremely often, then the sixth (or second box from the right) should be marked.

1 How frequently do you get emotionally involved (angry, sad, or happy) in the news stories that you read or hear?

|________|________|________|________|________|________|________|

NEVER OCCASIONALLY OFTEN

2 Do you easily become deeply involved in movies or TV dramas?

|________|________|________|________|________|________|________|

NEVER OCCASIONALLY OFTEN

3 Do you ever become so involved in a television program or book that people have problems getting your attention?

|________|________|________|________|________|________|________|

NEVER OCCASIONALLY OFTEN

4 How mentally alert do you feel at the present time?

|________|________|________|________|________|________|________|

NOT ALERT MODERATELY FULLY ALERT

5 Do you ever become so involved in a movie that you are not aware of things happening around you?

|________|________|________|________|________|________|________|

NEVER OCCASIONALLY OFTEN

---

[2] *Original version: Witmer, B.G. & Singer. M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225-240.

6 How frequently do you find yourself closely identifying with the characters in a story line?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

7 Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

8 How physically fit do you feel today?

|________|________|________|________|________|________|________|
NOT FIT MODERATELY FIT EXTREMELY FIT

9 How good are you at blocking out external distractions when you are involved in something?

|________|________|________|________|________|________|________|
NOT VERY GOOD SOMEWHAT GOOD VERY GOOD

10 When watching sports, do you ever become so involved in the game that you react as if you were one of the players?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

11 Do you ever become so involved in a daydream that you are not aware of things happening around you?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

12 Do you ever have dreams that are so real that you feel disoriented when you awake?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

13 When playing sports, do you become so involved in the game that you lose track of time?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

14 How well do you concentrate on enjoyable activities?

|________|________|________|________|________|________|________|
NOT AT ALL MODERATELY WELL VERY WELL

15 Have you ever gotten excited during a chase or fight scene on TV or in the movies?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

16 Have you ever gotten scared by something happening on a TV show or in a movie?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

17 Have you ever remained apprehensive or fearful long after watching a scary movie?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN

18 Do you ever become so involved in doing something that you lose all track of time?

|________|________|________|________|________|________|________|
NEVER OCCASIONALLY OFTEN